DOCUMENT RESUME

ED 057 004 SP 005 410

AUTHOR	Mood, Alexander M.; And Others
TITLE	How Teachers Make A Difference.
INSTITUTION	Office of Education (DHEW), Washington, D.C. Bureau of Educational Personnel Development.
PUB DATE	71
NOTE	173p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (Stock No. 1780-0813; $1.00)
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Educational Environment; *Individual Development; *Individual Differences; *Teacher Attitudes; *Teacher Behavior

ABSTRACT

The seven papers in this book were prepared as the bases of discussions at a BEPD-sponsored conference which dealt with the ways in which teachers can make a difference in education. The conference was a followup to one held a year earlier on the topic, "Do Teachers Make a Difference." The implications which can be drawn suggest the need to make American education more heterogeneous so that students can be offered a number of valid alternative choices. Teachers also need alternatives, as some function best in a structured setting while others perform more effectively with fewer constraints. Programs and the assignment of teachers must be tailored to the needs and aspirations of individual students, and teachers must be trained for this. The individual papers included are 1) "How Teachers Make a Difference," by Alexander M. Mood; 2) "The Difference Teachers Make," by Philip W. Jackson; 3) "A Tool-Development Strategy for Research on Teaching," by N. L. Gage; 4) "Structure and Teacher Performance: A Prologue to Systematic Research," by Dan C. Lortie; 5) "New Directions for Research on Teaching," by Barak Rosenshine; 6) "A National Coordinated Program of Research on Teaching Effectiveness," by Ned A. Flanders; and 7) "Learning Environments-or-Rooms for Thought," by Lawrence M. Stolurow. (MBM)

DISCRIMINATION PROHIBITED—

Title VI of the Civil Rights Act of 1964 states that, "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance." Therefore, Education Professions Development programs, like all programs or activities receiving financial assistance from the Department of Health, Education, and Welfare, must be operated in compliance with this law.

These papers were prepared by the authors for a conference sponsored by the Bureau of Educational Personnel Development, U.S. Office of Education. They reflect the findings and views of the authors and no official support or endorsement by the U.S. Office of Education is intended or should be inferred.

ED057004

OE-58044

HOW TEACHERS MAKE A DIFFERENCE

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SP 005410

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Bureau of Educational Personnel Development

Superintendent of Documents Catalog No. HE 5.258:58044

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1971

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402—Price $1
Stock No. 1780-0813

FOREWORD

In his Message on Education Reform in March 1970, President Nixon said, "We must stop pretending that we understand the mystery of the learning process, or that we are significantly applying science and technology to the techniques of teaching."

The papers which follow clearly substantiate the President's point. The state of research and development activities regarding the influence of teaching in the learning process is primitive.

These papers were prepared as the bases of discussions during a day-long conference in April 1971 at the Rayburn House Office Building. The conference, sponsored by the Office of Education's Bureau of Educational Personnel Development, was a followup to one held a year earlier on the topic, "Do Teachers Make A Difference?"

The earlier conference indicated that of all the factors that constitute a school, the single most influential in terms of pupil performance is the impact of the teacher.

The 1971 followup conference on how teachers make a difference was convened to assist the Office of Education and others responsible for formulating public policy in determining priorities in the recruitment, training, retraining, and utilization of educational personnel.

Obviously, the views expressed in the following papers are those of the authors and do not reflect official policies of the U.S. Office of Education.

However, certain policy implications can be inferred.

We need to make American education more heterogeneous so that we can offer students a choice of a number of reasonable alternatives. And this needs to be an honest choice where one alternative does not outrank another—where, for instance, students who opt for a highly structured educational experience do not get more credit or more status or more rewards than those who choose to embrace other styles of study in other settings.

If we are really going to individualize, if we are really going to humanize American education, we really can't do so unless we provide different approaches to meet the requirements of different people.

The same is true of teachers. Some function best in a structured setting, others perform more effectively with less constraints.

We must tailor our programs and assignment of teachers to the needs and aspirations of individual students; and we must do a far more effective job of preparing teachers to meet the individual needs of students.

Our goal is the development of more efficient and effective ways of delivering appropriate educational services to all students at all levels.

Finally, I wish to acknowledge the contributions of Dr. William L. Smith, acting associate commissioner for educational personnel development; of Mrs. Iris Garfield and Mrs. Patricia Wagner of the bureau staff; and of Mrs. Charlotte Hoffman of the Office of Education's public affairs staff.

Don Davies
Acting Deputy Commissioner for
Development

CONTENTS

Page

Foreword iii

Chapter 1
How Teachers Make A Difference 1
Alexander M. Mood

Chapter 2
The Difference Teachers Make 21
Philip W. Jackson

Chapter 3
A Tool-Development Strategy for Research on Teaching . . 32
N. L. Gage

Chapter 4
Structure and Teacher Performance: A Prologue to Systematic Research 51
Dan C. Lortie

Chapter 5
New Directions for Research on Teaching 66
Barak Rosenshine

Chapter 6
A National Coordinated Program of Research on Teaching Effectiveness 96
Ned A. Flanders

Chapter 7
Learning Environments—or—Rooms for Thought 123
Lawrence M. Stolurow

Chapter 1

HOW TEACHERS MAKE A DIFFERENCE

Alexander M. Mood

The immediate task of this volume is to assess the present state of the art of teaching and particularly to bring together in one place the views of the leading research workers as to how teaching can best be accomplished. The primary purposes in doing so are: (1) to give those responsible for the development of better teaching a current basis on which to plan their programs, (2) to give teachers themselves some information about how researchers believe teachers might improve their performance, and (3) to indicate where there are significant gaps in our knowledge about teaching and show researchers where they might profitably put their efforts. All of us concerned with better schools are much indebted to the six distinguished educational research workers who have prepared the ensuing chapters with great insight and care and honesty. I have learned much from them as will be evident when I try in this chapter to provide a general overview of where we stand with respect to teaching in the public schools.

Teaching Today

We have general knowledge about what comprises good teaching but we seem to be very far from understanding in detail how good teaching is carried out and hence we are unable to give explicit guidance to teachers regarding what they should actually do in the classroom to teach well. There are good teachers—superlative teachers. Philip Jackson gave us a rough guess that perhaps 10 percent of teachers are excellent; 10 percent are hopeless; and 80 percent are the masses who are doing the job with varying degrees of competence and conscientiousness, ranging from the masses who are goofing off part of the time all the way down to the hopeless who are plainly incompetent or who are shamelessly goofing off all the time.

There are several categories of excellent teachers. There is the commandingly authoritative teacher with high standards who accomplishes a great deal of learning by means of sheer personality power which does not permit students to neglect their studies. There is the teacher with an unusual store of infectious enthusiasm for everything in sight who is able without perceptible conscious effort to generate a great deal of enthusiasm for learning in most of the lucky students in hir[1] classes. Both of these categories include relatively rare people who would be successful at almost anything and it would not make a great deal of difference what kind of training they had or what kind of teaching methods they used; the students would learn in any case. Probably the most numerous category of good teachers consists of those who are very sympathetic to children and very concerned about their interests and welfare. They treat their students as people—not as inferior creatures prone to misbehave. The students are able to sense that this teacher is truly on their side and is plainly working for their benefit—as opposed to that teacher who claims that hir efforts are solely for their benefit when in fact some of those efforts are obviously designed to enhance hir own comfort or bolster hir own ego. The dedication of this teacher is enough to overcome significant deficiencies in hir knowledge or teaching techniques but clearly good training would be of a great deal more assistance to hir teaching than it would to the gifted teachers referred to earlier.

It is perfectly obvious that the bulk of teaching, by humans at any rate, must be done by ordinary mortals; we cannot expect the average teacher to become a gifted person or a saint either by training or by hir own efforts. Does research have any messages for those of us who are trying to do a professional job of teaching and who would give a fair trial to any reasonable-sounding research finding?

Yes, there are several messages, for example: *pay attention to what students say and put it to use.* No matter how trivial a student's question may be or how irrelevant to the matter at hand, seize upon it as a nugget of opportunity. Even if the question is stupid, treat it seriously and interpret it in such a way that an answer to it can make at least a small contribution to learning. If you cannot answer the question or can only partially answer it, say so and think of ways you and the class might ascertain the answer so that students will see that everyone, even knowledgeable teachers, must look things up and search things out. They will also learn how educated persons set about searching for information that is not readily available. Come back to the question next day, being careful to give further credit to the student who raised the fascinating issue in the first place. The object of this strategy is to give students a real sense of participation in the learning process, a sense that they have contributions to make, that they have thoughts about the process which have significance, that they have a measure of control over the process and hence can turn it to some degree toward things which interest them, that education regards their curiosity as important. Volunteered bits of information and suggestions by students should also be given the same deliberate attention. They should be elaborated upon; other students should be encouraged to elaborate upon them; you can bone up on them during the following day or two and elaborate further at an opportune moment.

Increase every student's sense of personal worth. On the purely negative side this means that students must never be given the impression that they are dumb or delinquent. Of course students will make mistakes but it must be made clear to them that everybody, including teachers, school administrators, and parents, make mistakes too and if these adults make fewer mistakes it is because they have learned by experience to avoid repeating certain kinds of errors and suffering their unpleasant consequences. On the positive side it is essential that a teacher find, for every student, some knowledge or skill or attitude that that student can be proud of and the teacher must go out of hir way now and then to compliment the student on that attribute.

Build up every student's sense of self-confidence especially with respect to his ability to learn. It is impossible to teach a child anything if hir confidence in hir ability to learn has been destroyed. Hence the greatest traitor to hir calling is the teacher who unwittingly erodes a child's self-confidence or gives a child the impression that he or she is dumb. The teacher can only praise progress and must be careful to take every opportunity to do so with slow learners. The teacher cannot be impatient with slowness but must take the attitude with the child that some people find some things difficult to learn and other people find other things difficult to learn, that there is no accounting for these differences, and that when we encounter them we must work especially hard at them as everyone else must.

The teacher must avoid assuming a posture of moral or intellectual superiority. For one reason, it is difficult to get away with; there are likely to be one or two students in the room who are brighter than the teacher, who will see through these poses, and who will be diligent about communicating their findings to their fellow students. That situation makes it very unlikely that the teacher will be able to catalyze much learning because hir credibility will be severely damaged. But even when the teacher can get away with them, these postures inhibit learning by requiring students to be continually on guard in order to protect their egos; they cannot therefore give full attention to what is supposed to be learned. Worse, these postures generate antagonism on the part of the students who will be inclined, as a matter of holding up their side of the conflict, to seek flaws in whatever the teacher says or does.

A teacher should analyze and attempt to minimize conflicts of interest with hir students. Or, to put it the other way, a teacher should try to enlarge hir community of interest with the students. There is no escaping large conflicts of interest. Children like freedom, whereas school is a stringent curtailment of freedom requiring daily attendance by law with the teacher serving as the jailor. Children like to play, whereas the teachers require them to work. Children are full of energy demanding to be expended in talk and movement, while teachers like quiet and order. On top of these built-in conflicts, there are various inevitable personality conflicts between teacher and a few unlovable students as well as certain behavior patterns on the part of some students which are obnoxious to the teacher. If the teacher is to facilitate learning, on the other hand, students must be convinced of hir good will toward them because learning requires the cooperation of the students and they will not be inclined to cooperate with an antagonist. A teacher should therefore go to some lengths to hide the conflicts by appearing to love every one of hir students and by permitting as much noise and activity in the classroom as possible so that it is not a terribly restrictive place for the students. Of course, noise and activity cannot be allowed to disrupt learning but it may be surprising how much students can tolerate and still learn; a teacher should try to wait for a student complaint before hir complaint is heard. The teacher should make clear that hir destiny is somewhat in the hands of the students because hir benefits are substantial in the eyes of the school administration when hir students learn well, so that there is a strong community of interest between their learning and hir career in addition to hir own personal satisfaction in the success of the students.

Some may object that there is something fundamentally dishonest about pretending great joy in one's work and one's students when in fact one often finds teaching to be humdrum and some of the students to be brats. The issue is simply resolved by keeping firmly in mind what the prime objective of teaching is. It is learning on the part of the students. It is not that one's role as a teacher be totally consistent with one's private life. A great many other issues and questions about how a teacher should do hir job can be resolved in exactly the same manner: by assessing the effects on student learning.

Minimize lecturing and passing down information from on high. One-sided instruction may seem to get the job done with greatest dispatch but it seldom does so in actuality. Learning requires effort on the part of the learner. If students are to learn something they must somehow be led to think about it—to participate in formulating the thing to be learned. Teachers must not simply inform but must raise a leading question; patiently let students struggle with answering the question out of their own experiences; and try to point out the unacceptable implications of answers that are not quite right, thus leading students to arrive at an acceptable formulation. It is a slow process, requiring great forbearance on the part of the teacher who will be sorely tempted to straighten out their muddy thinking at once and get on to all the other things that must be covered that day. But what do we mean by "covered"? That the teacher plowed through them or that the students learned them?

Maintain an open, relaxed atmosphere in the classroom. The object is to encourage participation by all students and every class has a few very sensitive students who find it terribly difficult to run any risk of embarrassment by opening their mouths. Hence there must be no risk of embarrassment; there must be no emotional penalty whatsoever for saying the wrong thing. Wrong ideas must be straightened out in a purely clinical way that focuses entirely on the way things are and the way logic operates and not on the mental processes of the contributor of the wrong idea. With shy students this must be done very gently, with great appreciation of their participation and with the exercise of ingenuity to find something right about the contribution that can be viewed with approval. Participation is essential for learning; a relaxed atmosphere is conducive to participation; a steady stream of approval and encouragement can maintain a relaxed atmosphere.

Create much activity for students. It is all right for them to learn to sit still but they should not actually do much of it in the classroom because to sit still for any length of time generates tensions in young people and tensions interfere with learning. Thus the school day should be rather thoroughly filled with things that require movement and talking and writing and drawing on the part of everybody in the class. Students can carry out all manner of experiments and constructions involving not only materials but plants and animals, perform dramas, do role playing, play intellectual and educational games, and practice skills and crafts. Small groups can carry out special projects simultaneously. Students can tutor other students. Much educational activity can be converted to contests between halves of the class, or between two or more teams with the remainder of the class serving as referees. If nothing better is available at a moment of restlessness, there can be a 5-minute break during which students can mill about the room and talk or sometimes do calisthenics or sometimes race 10 times around the room at top speed.

Use desperate means, if necessary, to combat apathy and boredom. It is very important that school not become a drag and an undesirable place that one seizes every opportunity to escape from. The reason is that the negative attitude about school gets transferred to learning; learning becomes something to be avoided. Teachers must be on the alert for boredom and have on hand at all times an arsenal of tactics for warding it off. If most of the class is interested in what is going on but a few are bored, those few must be separated out and put to doing something else that may be less educational but at least has the merit of not permitting antagonism to education to grow. If a sizable portion of a class gets bored it is definitely time to move on to the next activity even though the current one is not completed. If no educational activity seems to work on a lovely spring day, do something frivolous; take a hike; tune in a rock radio station and let the kids dance; dance yourself.

Show genuine concern for the well-being of every student in the class. Every student must be convinced by your attitude that you care about him or her as a person—not just about hir learning but about the whole person and hir whole life, health, comfort in and out of the classroom, family, successes, failures, and problems. Listen to idle chatter about these matters; give your personal help when they bring serious personal problems to you. A student will believe that what a teacher who cares is trying to teach must surely be worth learning, and will make an effort that will often put to shame all the clever teaching tricks in the book.

Do not let individuals fall behind. Identify early the difficulties that individual students have and move quickly with all the resources necessary to overcome them. You yourself may have to devote considerable time to a particularly difficult problem after school or perhaps even at the child's home so that the parents can be enlisted to help out. But most of the time less dramatic measures will suffice. Perhaps the student's classmates can form a team to help clear the hurdle; perhaps aides or counselors can help; perhaps older students in another class can be drafted as tutors. The thing is to treat it as a very serious matter—not just for that student but for the whole class. The whole class must pitch in to help even if it means slowing down their own progress so that the teacher can devote extra time to the student who is in danger of falling behind.

The object of this strategy is to convince all students that learning is extremely important and that everyone must learn; it is so important that it is unthinkable that a child should not learn.

Diversify your teaching. Be on guard against getting in a comfortable rut. Use all the tools of the trade available at your school—films, slides, audiotapes, records, curriculum packages, games, programed instructional devices, and so on. Use them professionally by becoming completely familiar with them and with how they are designed to be used before bringing them to the classroom. Read the journals to become acquainted with new teaching ideas and new tools of the trade as they become available; persuade your school to obtain promising new tools on trial and try them out in your class.

Consult your students about teaching. Let them in on the problem of making learning attractive and let them participate in solving it by presenting their own ideas as to how they learn best. Encourage them to help devise experiments which you and they can carry out to discover what seems to work well. Even if they are not very good experiments the Hawthorne effect will probably make them useful learning experiences.

Teach as well as you know how. While there is a great deal of encouragement for teachers to get additional education, additional degrees, and special kinds of training, many research workers find that teachers already know considerably more about how to teach than they put into practice. Some teachers find the use of certain aids too troublesome to bother with; they find that certain kinds of interesting projects require too much preparation on their part and hence decide, as a favor to themselves, that the projects are probably not as effective as advertised. The more teachers there are who refuse to cheat children this way the harder it will be, as a matter of social climate, for other teachers to cheat them.

Interact with your colleagues. Discuss teaching with your fellow teachers. Ask them how they deal with certain problems for which you have not found a satisfactory solution. Ask them to visit your class now and then in order to give you an appraisal of how well your class is learning and what might be done to improve it. Ask good teachers whether you may visit their classes in order to learn some of their practices. Spread the word among other teachers about your more successful teaching endeavors. Bring new ideas you have encountered in your reading to the attention of other teachers. Promote joint ventures with other teachers in which students of two or more classes will work together on some learning enterprise.

Do not participate in grading systems. Grades create an almost unbridgeable chasm between the teacher and those students who get the lower grades. Few things are so discouraging to a child as low grades; for many children nothing creates so much trouble at home as low grades. You, as the creator of that trouble, are well on the way to destroying your chance to help that child learn. Your position should be that there is a certain amount of material to be learned during the year, that every child is going to learn it, that you will certify to the fact that they have learned it at the end of the year, and that is that. This does not mean the material to be learned is necessarily the same for each child; an advanced or retarded child will likely continue to be advanced or retarded at the end of the year but the point is that you have set a reasonable learning goal for the child and have seen to it that the goal was accomplished. For those parents who demand to know where their child ranks it is quite enough for them to know the scores on standard achievement tests and thus rank their child against national norms. Every school should give such tests at the end of every year to every child in order to assess the school's performance with respect to every child.

Serve as a model. Of course that cannot be avoided. Day in, day out, students have before them the example of behavior of a very significant person in their lives—often the most significant person other than their parents, and sometimes even more significant than their parents. The example set may well be the most important thing a teacher does. It is hard to measure the effect of such an example; it is hard to observe even that it is having any effect; but research workers agree that the cumulative effect is tremendous and often changes lives. Personify intellectual honesty. Be enthusiastic about learning. Be humanly understanding. Share your life with your students. Show great tolerance in frustrating situations. Scrupulously hear both sides before trying to adjudicate any conflict. Explore unacceptable behavior with nothing but concern for the offender and how that behavior affects him. Know that your students are readily absorbing your skills, your knowledge, your attitudes, your value system, and your philosophy of life.

The Role of Administration

Can administrators help teachers make a difference? They can, indeed. The job of administration in any organization (school, business, or government) is to be continually aware of what is going on in the organization and to act when things are not going well. In school, therefore, the first and most important task of administrators is to function in such a way that when a student is not learning satisfactorily the matter will be brought immediately to their attention so that they can take corrective action.

Meanwhile administrators can create a climate in the school and policies for the operation of the school which will tend to minimize the occurrence of unsatisfactory events that will require them to take action. That is done by giving teachers and students every assistance to increase students' learning and bring about conditions which will make learning as easy as possible. Administration can, for example, do much to reduce conflicts between students and teachers by moving certain students out of classes of those teachers who find them especially difficult to deal with and also moving students from the classes of teachers who are unsympathetic to them. Much attention must be paid to grouping students and teachers in a manner that will maximize empathy throughout the school.

Administrators must give teachers regular attention and advice and appreciation. They should visit the teacher's class frequently and sometimes assist with the teaching in order to get a good feel for what is going on in the class. Only in this manner can they fully understand how they can best help the teacher facilitate learning. Only in this manner can they maintain familiarity with the day-to-day operations of the learning enterprise in their charge. Paperwork cannot possibly give administrators a comprehensive view of the delicate interactions between teachers and students that are so important to learning.

When administrators are visiting classes in order to help a teacher with the teaching they should do so with considerable preparation and planning. The object of their effort will be only secondarily to give the teacher relief. Primarily the effort will be a demonstration for the teacher of techniques that the teacher particularly needs to learn, and the demonstration should be so effective that there will be no doubt in the teacher's mind after seeing it that the technique must be added to hir repertoire.

Administrators must develop a very positive attitude toward their teachers. Teachers must be encouraged to bring their problems to administrators; when they do bring one in, administrators must immediately make it their own problem and, insofar as possible, take it out of the teacher's hands and deal with it themselves. This policy is to promote learning. A problem is something that interferes with learning. Removal of the problem from the teacher's responsibility and relieving hir mind of it enables the teacher to give that much more attention to the real business of the school—learning.

Implications for Research

The present state of our knowledge about teaching makes it clear that we have much to learn. Thus we find in the later chapters of this volume numerous proposals for extensive research programs. During the course of the conference at which the chapters were discussed a number of other critical matters were suggested that needed exploration.

For example, the point was raised, as it was in the conference that preceded this one,[2] that we desperately need instruments for measuring the affective achievement of students. There does exist the Loevinger sentence-completion instrument for measuring maturity. Beyond that we have nothing—nothing to measure sense of personal worth, self-confidence, social competence, or sense of responsibility. Teachers, parents, administrators, and the general public are agreed that these outcomes of education are as important as reading, writing, and arithmetic, for which we have endless measuring instruments. We do not have these badly needed instruments for the affective domain because they are much harder to devise, presumably. It was hard to go to the moon, too. Research must get to work on these instruments and produce at least a few crude ones rapidly.

Another matter that came up in the preceding conference was research in organization theory. There seems to be altogether too much acceptance on the part of everybody that the bulk of learning takes place now, and will forever in the future, take place when a teacher stands in front of 20 or 30 students. There are endless other arrangements of teachers and students possible and it is a good bet that some of them could be much more effective than the traditional one. I shall suggest a few that appear promising to me.

An organizational arrangement which permitted all students to teach, for frequent, brief periods of time, one or two or a few younger students would probably have several very significant advantages. Not the least of them would be the identification of each student with teaching and the better understanding students would have of what teachers are trying to do. The result should be a much greater feeling of community of interest between teachers and students.

The one-room school organization needs thorough exploration. This arrangement, instead of grouping students by grade or age, puts a number of grades and ages in one room. The two obvious ways to do this are to: (1) have each teacher select from the student body of the whole school a group that would be most desired for hir class, and (2) let each student select the teacher of hir choice. There might be some problems of class size in the second case, in that some teachers would find themselves with very large classes and others with no students at all; but these problems might not be too serious. In large classes it would be expected as a matter of course that the older students would do considerable instructing of the younger ones; perhaps some of the teachers with no students would go away quietly and permit the administration to employ persons more facilitative of learning; those who did not go away could become assistants to the teachers with very large classes—perhaps in ways that would not entail a great deal of contact with students.

Of course there are other ways that one-room schools could be formed. There might be arbitrary assignment of students to teachers by a team of psychologists who had interviewed teachers to form some sort of personality assessment of them and who then interviewed students to judge which of them had personality characteristics which would mesh well with those of a given teacher. It would also be worth trying psychological instruments as devices for matching students with teachers. Another mechanism would be to let students make choices by seeing some of the other students, as well as the teacher, in a given class. Perhaps the oldest students would attach themselves to teachers first, then the next age group could make selections from among the nuclei thus formed, and so on down the age levels. The process could be reversed with the youngest students forming the initial nuclei and the teachers joining them last. Another version would permit any subset of persons to form a nucleus and let such nuclei grow by accretion to some size limit.

Another mode of organization would try to take full advantage of class activities that require little teacher participation (quizzes, written work of any kind, certain kinds of art work in which each student is fully occupied in hir own individual activity, etc.) to form large groups of students which could be supervised by one teacher and thus create time for other teachers to work with small groups of students needing intensive attention.

Another organizational mode would attempt to ascertain the greatest skill or skills of each teacher and then construct a schedule of classes and teaching that would, as nearly as possible, have each teacher spend full time working in the areas of hir greatest skill.

There needs to be extensive research on school organization appropriate for the new computer-related technologies that will soon be making their way into the schools. The present belief of research workers is that these technologies will be mainly appropriate for the cognitive domain and that students will pursue their studies with the new technologies very much on their own without close supervision of a teacher. Thus teachers may become primarily occupied with the affective domain and it is most likely that very different forms of organization from the present one will be most suitable for that kind of learning. In particular, the units of the organization should doubtless cut across age levels for many purposes.

Incidentally, we may notice that the affective domain is now in considerable need of curriculum research and that the need will become much more acute in the future as new technologies move into the schools. The reason is that students will be spending more time in isolation with machinery; hence the time they spend in the society of others will become increasingly important to their affective development and therefore must be used more effectively.

Another very important area crying for research is the matter of incentives for both students and teachers. For many, many years parents have used money and privileges as incentives for their children to work for good grades. Recently a few schools have tried out money and stamps exchangeable for prizes as incentives. Schools have long rewarded good students with recognition in the form of honor rolls and special privileges at school. However, we know very little about the effectiveness of these incentive devices; we are far from being able to tell school administrators how to construct systems of incentives that will appeal to different categories of students or to give administrators estimates of how much various incentives will increase learning.

There is also the question of incentives for teachers. It is generally agreed that there is no incentive in the perfectly flat salary schedule which pays the best teacher in a district exactly the same salary as the worst teacher who happens to have the same education and experience. Teachers get nothing extra in their paychecks for teaching well. (That is not quite true because districts with higher salary schedules seek out and make offers to the better teachers in lower paying districts; thus a good teacher can get a financial reward if hir family situation permits moving to another district). There are precious few other rewards, either, for good teaching; sometimes it is not even recognized by fellow teachers or administrators so that it does not even bring a pat on the back. School boards and administrators sorely need solid estimates of how much learning might be increased by salary and bonus incentives for teachers. There is also an urgent requirement for adjustment procedures which will equitably allocate financial incentives between teachers with various categories of students; that is, an increment of learning after 1 year of teaching students with low ability or preparation would be worth much more than the same increment accomplished with students of high ability.

Some extremely valuable experiments in school operating policies are taking place in various kinds of free schools which give students various degrees of control over the curriculum. So far as I know, few if any research workers are taking advantage of these experiments to discover how they affect learning and motivation. It would be most valuable to develop information regarding what categories of students learn better in free environments and what categories learn better in the more controlled environment of the traditional school.

Along the same line, we need research to investigate ways of operating a controlled and a free environment in the same school. Almost every school has students who are seriously restrained and frustrated by the controlled environment and hence learn little. To what extent might their learning improve if they could transfer to a segment of the same school operating in a freer mode? How free should that part of the school be to accommodate the majority of the frustrated? Would those in the free segment learn more or less in the affective domain? If some learn considerably less in certain areas of the cognitive domain, should they lose their option to attend the free segment? Or should the whole matter of the transfer be the choice of the parents?

What would be the consequences for school operations if parents were given the option to specify that certain aspects of learning in both the cognitive and affective domains be emphasized or reduced for their children? What would be the consequences for learning?

I have tried to raise here some crucial questions that can only be answered by competent research projects and which must be answered if we are to design and develop a more responsive system of public education. My general impression about educational research is that it is too much inclined to take the existing system as given and to search for ways of improving the system. An exception to that statement is the work of those who are investigating computers on the assumption that computers will bring about a substantial revolution in education, will completely alter the organization and operation of the school system. There are a great many other ways to bring about a revolution, though, and research should be testing some of those in the hope of uncovering some promising answers—there must be real hope for improvement in consideration of the fact that we are dealing with an ancient form of organization and mode of operation. The odds are that focusing on the existing system or assuming that the existing system will remain largely intact will put an unrealistically low ceiling on what might be achieved.

Finally, in commenting on research, I wonder how promising it may be to explore in great detail the actual teaching activity in the classroom by recording bits of behavior on the part of teachers and students. When we do that we eventually correlate these bits with student achievement and expect to discover certain patterns of behavior that are associated with high increments in achievement. The research could very well give us the wrong answer because high achievement is most associated with exceptional teachers—teachers who are inspiring or enthusiastic or dedicated or have strong personalities. Their methods may not be very good; their classroom behavior may not be at all effective for the ordinary teacher. There is the further difficulty that teachers are very different and have very different talents so that behaviors appropriate for one person will be quite inappropriate for another person. Teachers have to find their own styles by cut and try; there is much that we can do to help them by educating them and showing them a variety of teachers in action but we may not be able to analyze a teacher's personality and use that as a basis for prescribing hir classroom behavior patterns in detail. I would bet, at any rate, that a rational system of teacher incentives would be much more productive of learning than detailed prescriptions of teacher behavior.

Implications for Development

The present state of our knowledge about teaching seems to justify the following very far-reaching conclusion: Teaching is difficult and complicated; only a genius can do an exceptionally good job of it; if we are going to achieve a distinctly higher quality of teaching we must either transform teachers into geniuses or we must simplify their job by providing them an extensive array of sophisticated tools and equipment. That conclusion leaves the field pretty wide open for development.

Perhaps we can get one clue to the future of teaching by observing a development in medicine being tried out by the University of Minnesota. Its Medical School operates a large computer which has cataloged in its memory a fairly complete list of human ailments, together with the complexes of symptoms associated with each, the optimal treatment, and all the cautions that must be observed with respect to the treatment and how it must be modified for exceptional cases or exceptional patients. General practitioners in Minnesota can rent a small console for their offices so that they can consult the computer over a regular telephone line. Thus a doctor can use the computer as a diagnostician by giving it a list of the patient's symptoms and the computer will search through its memory to find one or more ailments which fit the symptoms; if there is more than one, the computer will tell the doctor what additional information must be obtained in order to identify the ailment uniquely. Then, if the doctor requests it, the computer will transmit the preferred treatment for the ailment according to latest research information. Doctors love this system—especially older ones who have been out of school for some time and have been too busy to keep up well with the medical literature; they were formerly accustomed to worry that they might not be giving their patients the most beneficial treatment in some cases; the computer not only erases that worry but gives them the confidence of being backed up by faculty of the University of Minnesota Medical School. It also erases the necessity to read the literature and waste a lot of time learning about new developments that one may never encounter in hir practice (that is why doctors often let their reading slip); this is a very large benefit to doctors who are normally extremely busy. One computer can serve thousands of doctors.

This medical consulting system may well be the forerunner of the first sensible educational system. The human body is quite a complicated apparatus but at least it tends to obey the laws of biology; the human mind is, in many respects, still more complicated and that is what teachers are trying to deal with. They are doing it with much less training than doctors have, with vastly less adequate research information on which to base their work, with far fewer tools, with essentially no specific prescriptions at all, and with batch processing. Clearly the whole business is ridiculous and it is a wonder that it accomplishes anything at all.

So we say to development: Start building a sensible educational system which will turn teachers into doctors. Provide them a complete array of sophisticated diagnostic tools. Create fascinating learning tools which students can use, largely unassisted, to correct any sort of educational deficiency. Professor Gage has listed several types—audiotapes, videotapes, programed texts, computer-assisted instruction, kits, manuals, models, simulators; to that list we may add games—hundreds of specially designed games focused on specific educational aims—and newspapers, libraries, museums, concert halls, theaters, factories, farms, courtrooms, offices, stores, parks, toys, motorcycles, musical instruments, crafts, pets, vegetable gardens, and so on. (A full year's physics course can be hung on a motorcycle; it has everything—statics, dynamics, heat, light, sound, electricity, magnetism, optics, power, combustion, mechanics, energy, friction, thermodynamics, stresses, and equilibria.) Of course the diagnostic side of the educational system will be very different from that of the medical system because the computer will keep continually on call a complete record of a child's educational progress in both the affective and cognitive spheres and will be ready to make recommendations each time evidence of a step in the child's educational progress is added to the record. It will be quite a nice development problem to determine what kind of evidence will best serve as a basis for sound recommendations.

Implications for School Boards and Superintendents

Since research has far to go in order to determine how education should be conducted and since all these fancy computer developments are obviously years away, perhaps schools should just rock along in their present modes and await developments? No! If education is to be substantially improved, schools themselves will have to carry out much of the required development.

One reason lies in the fact that development is very expensive and there are not available at present nearly sufficient funds specifically earmarked for development to do even a small fraction of what must be done to make significant progress. Another reason lies in the fact that even the little money available will tend to go to the more exciting and showy hardware. Of course school organization and policy and personnel are more important than the hardware (procedures for its use may be more important than the hardware itself), but persons allocating development funds are notorious for enjoying a beautiful shiny new touchable gadget at the end of the line; it seems to give them assurance that the money really did buy something. A third reason is that these less spectacular developments cannot be carried out very well outside a school; a fair trial of them requires participation of teachers and students in a real educational situation and only schools have that resource. The fourth and most compelling reason is that other agencies may not give due weight to the desires and needs of schools. Universities tend to put a great deal of emphasis on buttressing academic reputation; commercial firms tend to put a great deal of stress on the generation of future profit; the pressing interests of schools may not always be best served by these tendencies, which have not in the past demonstrated great community of interest with school needs and desires. A fifth reason is that these other agencies do tend to look closely at what goes on in schools and will take an interest in whatever innovative programs they detect; thus schools can, by going directly to work on their pressing problems, exert considerable influence on what other agencies do.

The most important thing schools can do along this line is to be wildly receptive to any idea that anybody wanders in with for doing things differently. That is not easy to do. Top officials got there by being unusually familiar with all the written and unwritten rules and understandings inside and surrounding the organization. Any highly novel idea is almost certain to clash with no small number of these rules and understandings. Furthermore, the perpetrator of the idea is likely to be unaware of these clashes or to consider them fairly unimportant. Nevertheless, I am suggesting that officials rise up and embrace this nut and hir nutty idea. It won't be easy. By *embrace* I mean go all out to arrange a trial of the idea that will enable the proposer to see its consequences in practice. It is all right to point out to the proposer how the idea violates the State education code and various district regulations, provided one suggests at the same time ways around these difficulties and later endeavors to secure waivers of regulations as required. The aim is to give the idea a test, preferably under the supervision of the proposer, if it is at all possible. Even if the idea is the epitome of idiocy, there will be advantages to giving it at least a small pilot trial: (1) teachers, students, and parents will get the message that this school is commendably concerned about improving its operations and open to suggestions—that posture might bring in a good idea now and then; (2) the morale of the proposer will take a big jump and hir work should, as a result, become much more effective; (3) those working with the proposer on the idea will have their own morale improved by the association with a person who is unmistakably enthusiastic about hir project. And of course there is always the Hawthorne effect to prevent the trial from being a complete waste of effort. There is even a slim chance that the trial may confound all expectations and show the idea to have some merit. On a rare occasion one will hit the jackpot and generate a real improvement which will jump from district to district clear across the land and it will probably have looked as "way-out" as any of them when it was first proposed.

The other very important thing that officials can do is generate their own ideas for reforms. Members of the school board, especially, must not be bashful about promoting their own proposals because they believe educators know a lot more about school operations than they do. Educators do know a great deal about an ancient obsolete system that is still wobbling along but has not much longer to go. The question is: What will replace it? Educators are so immersed in the present system that they are less likely than anybody to be able to conceive what may replace it. So others must start developing those conceptions and surely school board members head that list because they have prime responsibility for the schools. They have an overall view and understanding of school operations; they have experience with efficient practices in other kinds of organizations—some of which can be transferred very profitably to the educational organization; they can make good judgments about which ones might work because of their knowledge of practices both inside and outside schools; and they as representatives of the public can best formulate and keep their eyes on the desires of parents for their children.

One innovation with great potential, for which a good case can be made at present, is a system of educational accountability. There is widespread support of the idea on the part of many distinguished educators so that it will not be easy for a local educational organization to resist the insistence of the board that such a system be adopted. The fundamental object of the system is to get measures of how much learning is taking place in the school by testing every student at the end of every year and comparing hir test scores with scores for the previous year to see whether hir advancement during the year has amounted to one grade level or not. Of course there will be considerable variation in the actual increments; some will be for more than one grade level and some for less. The board needs to understand the reasons why; some will be good reasons, some will be bad; something needs to be done about the bad reasons. And that, of course, is the point of the accountability system—to improve the learning of children by doing something about the things that are inhibiting learning or not pulling their weight in increasing learning. Sometimes those things will be teachers. School boards must examine the record of each teacher in raising the grade level of hir charges. A teacher who does not seem to have done well may have had many difficult slow learners in hir classes and may in fact have done a wonderful job. Therefore hir record for the year must be judged in the light of what other teachers did in the preceding year with those same students. By the same token, every teacher's record must be compared with the record made by hir students in the preceding year under other teachers; a teacher who raised every student more than one grade level may, nevertheless, have done a poor job when one looks at what other teachers accomplished with the same students the year before.

This accountability system will often be resisted by teachers because they naturally will not relish being put on the spot when learning seems not to be going well. Since educational administrators were usually once teachers, they will tend to sympathize with teachers and will sometimes resist it also. One reason that educators at a higher level endorse accountability is that they are looking further ahead to the day when there will come into being the elaborate computer systems and audiovisual systems which are slowly being developed. When that day comes they want a good evaluation system in place and operating in local school districts so that school boards can judge for themselves on a sound basis whether such and such a beguiling hardware system is really worth the money in terms of student learning. For this reason it is absolutely essential that the system of accountability include tests in the affective domain as well as the cognitive domain. It is true that we now have only primitive tests of this sort but they are better than nothing and should be used. Once testmakers see that there is a market for these they will undertake to develop some decent tests. Hardware systems are judged now by many educators to be well suited for many kinds of cognitive learning but not useful for most personal development and development of interpersonal skills. We must get some kind of solid grip on these affective aspects of learning so that they will not be neglected when the hardware comes in.

The system of accountability is a fundamental innovation that needs to come early to any school system because it is a valuable tool for assessing the effectiveness of all other innovations. That is, long before computer and other hardware systems arrive, the district may be trying out all kinds of teaching and administrative and organizational ideas and will be in a much better position to evaluate them if it has a system of educational accountability in operation. Thus the system will serve officials in two ways: by pointing out where improvements need to be made and by helping them to judge whether alleged improvements are performing as claimed.

While I have been at some pains to emphasize the importance of such a system I must say also that it has its limitations. There are all kinds of extenuating circumstances that make numbers come out the way they do and, in making decisions based on those numbers, decisionmakers must hear and give due weight to them, so that the decisions will reflect good judgment about all the factors that the numbers do not cover and all the unexpected circumstances that diminish the validity of the numbers. The system may be likened to a firm's accounting system. The dollars-and-cents story of the operations of various segments of a firm provides valuable information for the firm's executives but it is not the whole story by any means. Decisions based on financial data must be tempered by consideration of the general state of the economy, the stringency of the competition in certain sectors, random external perturbations that could not have been foreseen, special demands on certain segments that reduced their profitability but benefited the firm as a whole, and so on. On the other hand, executives would be crazy to try to run a firm without an accounting system.

Footnotes

[1] *Hir,* pronounced "here," is used for "his or her."

[2] See *Do Teachers Make a Difference?* U.S. Government Printing Office, Washington, D.C., 1970, OE-58042.

Chapter 2

THE DIFFERENCE TEACHERS MAKE

Philip W. Jackson

> "... it is irritating to find oneself in the presence of something that is both important and indefinable."
>
> Eric Newton, *The Meaning of Beauty*

* * * * *

This essay is about the effects teachers have on their students. It begins with an effort to replace an overly simplified view of what teachers do with one that is more complex and, hopefully, closer to educational reality. It ends with some speculations on how teaching, as complexly perceived, might be improved. The perspective from which it is written is that of an educational researcher turned practitioner, an academic who of late has become increasingly involved in the practical affairs of schooling.

Education, as every schoolboy knows, has to do with the acquisition of knowledge and skills. From kindergarten onward schools are places in which the ignorant are expected to become more knowledgeable and the maladroit more skillful. The chief business of teachers, as every schoolboy also knows, is to assist in these transformations. On that much there seems to be perfect agreement. Beyond it lie educational controversy and confusion.

Part of the controversy, if not the confusion, arises from our efforts to speak about knowledge and skills in the aggregate and to make comparative statements about these abstract assemblages. Because there is so much to know, from letters of the alphabet to Einstein's theory, and so many skills that might possibly be mastered, from operating a zipper to performing heart surgery, we often find ourselves in the position of wanting to say something about the total amount of knowledge or skill a person or group of people possesses. Thus, we speak of parents as knowing more than do their children about most things; we describe ourselves as having learned a lot or a little from an experience; and we are quick to acknowledge that some people are very skillful, while others are not.

These ways of talking in crude quantitative terms about what people know and can do create little difficulty in our everyday conversations. When someone says, "She knows more than he," we all understand the message; its form is intuitively satisfying. Occasionally we may wonder whether or not such a statement is true and, consequently, we may question the speaker's evidence. But never do we question the legitimacy of applying the idea of differing amounts and degrees to our descriptions of human achievement. Rarely, if ever, does the average person pause to consider the set of assumptions upon which such judgments rest. He is even less likely to worry about how the precision of such descriptions might be improved.

In such matters the educator differs sharply from the man-in-the-street. Making statements about how much people know and can do is central to his professional role. As a result, the application of quantitative concepts to such descriptions has developed, within educational circles, into what might almost be called a fine art. The educational testing movement, which has become tremendously influential within the past few decades, stands as a massive and systematic attempt to bring greater precision to the statements professionals make about a wide range of individual differences—differences about which the average layman is quite content to be vague. In today's schools the chief instrument for the attainment of that greater precision is, as we all know, the paper-and-pencil test of academic achievement.

The enormous amount of money and human energy spent in the development of these testing procedures is perfectly understandable when we consider the range of their applicability to the problems educators face. Miss Jones senses that Billy knows more than does John about the geography of South America. But does he really? And, if so, how much more? The developers of some educational materials have a hunch that students will master the content faster if it is presented in one way than in another. But is this true? And, if so, is the difference sufficiently great to bother about? The answers to questions such as these—and the list could easily be extended to any desired length—are customarily sought in statistical summations of students' responses to paper-and-pencil tests of academic achievement.

Given what has already been said about the teacher's role in assisting students in the acquisition of knowledge and skills, it seems only natural to extend our quantitative approach in that direction as well; only natural, in other words, to search for answers to questions of teacher effectiveness in the scores derived from tests of one sort or another. Thus, we describe Mr. Brown as a good teacher because his students give evidence of having learned a lot in his class. Brown's colleague, Mr. Smith, is thought of as having done poorly because his students have learned little. How do we know this? Well, there for all to see are the class averages on the year-end tests of achievement.

On the face of it, such a procedure for comparing the quality of two teachers makes a lot of sense; so much so, in fact, that some educators suggest applying it to all manner of judgments about teachers and the effects they have on their students. For example, some would insist that teachers be held accountable for a "year's growth" in all their students as measured by test performance. To this, others would add the requirement that teachers' salaries be made to reflect differences in achievement "gains" as so measured. The latter suggestion calls to mind the old "payment by result" practice that was used in England at the turn of the century.

Attractive as they might appear at first glance, such proposals have all sorts of hidden difficulties connected with them. For example, there are any number of technical problems that would have to be solved before such proposals could be put into practice, including such matters as the adequacy of the measuring instruments, their validity and reliability, and methods for dealing with different learning rates and different backgrounds of experience. Far more important than such technical matters, however, are the conceptual shortcomings of this way of looking at teacher effectiveness. These shortcomings have less to do with what tests tell us about a teacher's success or failure than with what such tests do not tell. In other words, the crucial question is: What are we likely to overlook when we adopt the "achievement gains" point of view?

To adults who have never taught and have seldom pondered the complexity of teaching it must seem as though people who manage classroom affairs could be best described as master show-and-tell artists. Even students, unless they are very discerning, are likely to share this perspective. The teacher's job, at least as seen from the outside, would seem to consist chiefly of standing before groups of students and telling them what they should know or showing them how to do things. It is true that students are commonly called upon to do some showing and telling of their own, which the teacher then evaluates, but the onus of responsibility for the success of the activity is on the teacher. That success depends in large measure on his talent as a conveyor of information and a demonstrator of skills; or so it would appear to those for whom the ends of education were reducible to how much students know or how well they perform a task at the termination of their instruction.

But, as a moment's reflection reveals, the experience of going to school leaves deeper marks on us all than those reflected in the surface scratches of testable knowledge, however important the latter may be. It is the teacher's contribution to those less visible marks that must be pondered. To do so, however, we must move beyond the stereotyped image of the teacher as a show-and-tell specialist, acknowledging, as we do, that such an image is by no means entirely false. Teachers do indeed engage in their share of showing and telling at all levels of schooling, but, depending on the age of their students, the particular circumstances of their work, and their individual talents and temperaments, they invariably do much more.

Ideally, teachers influence as well as instruct. As persons they have an impact on their students that transcends the content of lessons and assignments and is not immediately observable in what is said and done in class. They not only pass *on* knowledge, they *pass* on knowledge, judging it, criticizing it, imbuing it with an importance and a degree of human significance it would not otherwise have. In the face of all that could be known the business of deciding what knowledge is most worth having is of paramount importance. Teachers aid in making such decisions. Through their efforts students are made aware of their ignorance but, if properly handled, they are also protected from being overwhelmed by that awareness.

In the process of showing students how to do things teachers demonstrate that such things can be done. Through personal example, and customarily without fanfare, they make the idea of mastery credible and its attainment close at hand. They are, or should be, useful to students not simply as storehouses of information or as skillful performers (for books and films will do those jobs as well) but as model knowers and doers whose physical presence across the room brings the knowable and do-able almost within reach. Moreover, they are available not only as models to be seen and heard, but as viewers and listeners—as persons who are responsive to the words and actions of those who would seek to be like them. Thus, the full potential of the teacher's influence includes heightening the significance of learning in all of its forms through the fact of his personal and visible involvement in it.

With very young students there is additional scope to the influence teachers might have. For most of those students classrooms are the settings in which impersonal authority is first encountered. There, in the company of his age peers, the child learns to make his way in a network of dos and don'ts and to operate under a set of implicit and explicit standards that apply equally to all. As he helps students come to grips with these facts of institutional life, the teacher becomes an arbiter from whom important notions of fairness and of social justice are first derived.

Often the teacher is also the first adult outside the family to express genuine concern for the child's well-being and to care demonstrably about his progress or lack of it. Teachers who succeed in communicating that concern are surely contributing something to the education of their students that can never be supplied through any other means. Whether or not such a contribution is ever reflected in a heightened level of achievement we probably will never know, but to imagine that such is the only measure of its worth is to labor under an unnecessarily limited notion of what education can and should be about.

With older students teachers have the additional opportunity of personifying the virtue of intellectual honesty and the value of impartially seeking the truth. Perhaps few of them succeed in accomplishing such a noble end, but those who do often become more memorable than the facts they teach. Even those who fail may leave a residue of influence that lingers long after the dismissal bell rings.

This brief description of the impact teachers might have upon their students is admittedly idealized and hardly more than suggestive, yet it should suffice as a reminder of the many different ways in which teachers make a difference. Though they do indeed assist in the acquisition of knowledge and skills, that assistance is by no means limited to straightforward efforts to increase the amount that is known and the proficiency of the skills that are being taught. In addition to being direct and straightforward, teachers also work in ways that are indirect and subtle.

The subtlety and indirectness of the teacher's influence create several problems for teachers themselves and for those who would help them at their task. There is, first, the unfortunate fact that the conditions many teachers would most like to bring into being lie outside their control. What teacher, for example, would not aspire to increasing the love of learning among his students? The problem, of course, is that no one knows quite how to achieve that desirable state of affairs. Part of the difficulty arises from the vagueness of such a goal, but the recommended practice of reducing it to a set of molecular behavioral objectives only succeeds in making the whole effort ludicrous. As a result, the teacher who is concerned about such matters has no alternative but to behave sensibly and hope for the best.

A second problem, which is closely related to the general unpredictability of the teacher's success, is created by the equally disturbing fact that not all students respond in the same way to identical treatment. There is, indeed, a phenomenon which some psychologists today refer to as "attribute-treatment interaction," but their discussion of it does not begin to reveal the complexity that is part of the intuitive knowledge of every classroom teacher. As even the novice soon learns, the lesson that fires the imagination of some students leaves others cold; the teacher who is perceived by some as a paragon of virtue is seen by others as a pompous old fool. In education, as in life, one man's meat is another man's poison.

Third, even if a teacher is successful in transforming the process of learning into something that is exciting and vital in the lives of many or most of his students, there is no guarantee that his success will be perceived as such by outsiders. Even his students may not understand what is happening to them and may fail to give him the credit he is due. Though teaching is not without its rewards it is also, in many ways, a thankless task.

The chancey quality of the teacher's indirect influence would seem to provide grounds enough for forgetting the whole thing. Why not simply concentrate on the measurable effects of instruction and leave the rest to Lady Luck? Such an evasive tactic might be acceptable were it not for the fact that many of our most crucial educational problems have less to do with the mechanics of teaching a greater amount in a shorter time than they do with such intangibles as the value placed upon learning and the importance of schooling as perceived by vast numbers of today's students.

Anyone who has talked to teachers about their work or who himself has spent much time in classrooms knows that the roots of our educational ailments do not lie in the fact that we are not teaching students *enough*. Though the amount of their learning may not be as great as we would like, such a condition is merely a symptom of a difficulty whose source lies elsewhere. That difficulty, if we accept the testimony of practitioners and believe what we see in classrooms, has more to do with the phenomena of student apathy, disinterest, and boredom than with the mechanics of how learning takes place. It is unfair to hold teachers responsible for these many motivational ailments but they cannot be freed of their responsibility for doing their best to correct them. To repeat, such responsibility requires that they seek to influence as well as to instruct. Indeed, they likely will do so, for better or for worse, whether they seek to or not.

Though the potential influence teachers might have on their students is an important topic for discussion, it is also a subject that can be easily spoiled by sentimentality. Whenever we speak of how powerful his influence might be we run the risk of idealizing the image of the teacher to a point where it begins to resemble the hero in a Victorian novel more than any real-life teachers we know. Perhaps the bonds of the believable have already been exceeded in the description presented thus far. Surely it is true that few teachers, if any, come close to realizing the full range of positive influence that has been described. Indeed, if we could ever assess the totality of their impact, we might well discover that many teachers do more harm than good. As one of my colleagues is fond of reminding me: if classroom teachers are to be likened to artists, the sad fact is that the few Picassos and Cézannes among them are far outnumbered by the sign painters!

Yet even if we acknowledge that the talents of the average teacher leave much to be desired, an idealized image of what his influence might be is useful as we consider the way in which those talents might be enlarged upon. If most teachers have a very limited beneficial, or even a detrimental, effect on the way their students view the process of education, that is a problem to be solved, not a condition to be taken for granted. It is to that problem or, as we shall see, that set of problems that we now turn.

Any discussion of how teaching might be improved must distinguish between efforts involving altered working conditions and ancillary aids—such as smaller classes, the employment of paraprofessionals, better curriculum materials—and those aimed at improving the teacher himself. Both sorts of improvement are doubtlessly necessary. Moreover, it is uncertain as to which will result in the greatest overall good in the long run. Perhaps a reduction in class size or the development of better textbooks would do as much to enhance the quality of education for most students as would a comparable amount of energy spent in trying to increase the effectiveness of teachers through focusing on the way they work. However, it is only the latter with which we are concerned at present. Accordingly, all other schemes for making our schools better will be ignored.

One of the first questions to be faced by anyone setting out to improve education through work with teachers is whether our present educational shortcomings derive from a lack of knowledge or whether their source lies elsewhere. In other words, is the reason teachers are not now doing better than they are because they do not know how, or do other hindrances—e.g., personal apathy, economic limitations, tradition—stand in their way? Moreover, if our present shortcomings, or part of them, stem from a lack of knowledge, is that lack an ignorance suffered by all (such as our ignorance of a cure for cancer) or is it rather an instance of knowledge that is not widely enough shared (such as the secrets of a master chef)? Depending on our answer to this question, or our hunch as to the direction in which the answer might lie, quite different strategies for the improvement of education emerge. If, on the one hand, we believe that significant progress awaits the discovery of knowledge that no one yet possesses, we obviously would want to initiate the quest for that knowledge with all deliberate speed. If, on the other hand, we believe that superior ways of teaching are already being practiced in some classrooms, as compared with others, then the problem becomes one of ensuring that extant knowledge is more widely disseminated and put into practice.

As an educational researcher turned practitioner, my opinion is uncomfortably split on this question. My academic training and research experience pull me in one direction; my more recent experience as a school administrator pulls me in the other. The researcher in me prompts the desire to insist on the fundamental importance of seeking new knowledge about how to teach—and preferably knowledge that is derived from controlled empirical study. That part of me clings to the belief that bigger and better research projects (more subjects, more sophisticated data analysis, generously funded by the U.S. Office of Education!) will ultimately succeed in uncovering the secrets of teaching. Yet I confess that such a strategy has not been eminently successful to date. Moreover, I privately doubt that the results of future studies are likely to have a revolutionary impact on the practice of teaching. In short, I do not envision any educational researcher suddenly bursting from his laboratory (or computation center) and running into the streets shouting, "Eureka!" This is certainly not to say that educational research should be discouraged or that it cannot be of direct benefit to practitioners. The point is merely that we should not be unrealistic in the hopes we invest in such ventures.

Fortunately for those of us who must go about the daily business of keeping schools running, my more recent experience as a practitioner sustains the hope that a significant improvement of teaching need not wait upon the results of yet another empirical study. As I have talked to teachers, watched them at their work, and thought about the complexities of their task, I have more than once recalled the story of the newly graduated agronomist who was trying unsuccessfully to convince an old farmer to make use of new scientific methods in the running of his farm. Seeing that his appeals were falling on deaf ears, the agronomist finally asked in desperation, "What's the matter with you, Mister? Don't you want to improve on what you are doing?" "Shucks, son," replied the farmer, "I ain't doing as well now as I know how!"

If they were completely candid I suspect that many classroom teachers would echo that old farmer's confession. They are aware that they are not presently giving their work the full measure of their interest and energy. They are also aware that their teaching would be better if they did. For them the problem is one of reawakening their sense of professional pride and responsibility.

A second source of our present educational shortcomings seems to derive from the fact that teaching is such an engrossing activity that many teachers do not have time to look at their own practice and that of their colleagues with a critical eye. Moreover, even when they are given the opportunity, many of them do not know what to look for. In other words, they lack a critical stance from which to examine the process of teaching in general and their own work in particular.

Because many teachers have relatively little opportunity to see their colleagues at work, techniques and tricks of the trade tend not to be easily communicated and shared. As a result, each teacher becomes, in a sense, a solitary craftsman, inventing on his own hook procedures and ways of working that have already been perfected by his colleague down the hall. Though some of this craft and lore is picked up during training and more is passed on through chance conversation in the teacher's lounge, much of it remains locked within the teacher's "shop"—his classroom. This state of affairs, which is particularly unfortunate for the beginning teacher who has not yet worked out his own way of doing things, suggests yet a third avenue along which to pursue the improvement of teaching.

Here, then, are three sets of problems in need of solution: how to encourage teachers to become fully committed to their professional role—to try harder, so to speak; how to help them look critically at their own work and that of their colleagues; and how to make available to them the practical knowledge that is already possessed by others. From a practitioner's point of view an attack on these problems promises to yield rich dividends. The remainder of this essay is devoted to some speculation on how such an attack might proceed.

The fact that some teachers are only half-hearted in the performance of their duties is apparent to anyone who has spent much time in a school of any size, even, I might add, if it is a school that tends to attract teachers whose training has been above average in quality and whose professional commitment is reputed to be high. Though it is difficult to estimate the seriousness of this problem, I suspect that few students escape being taught by such a teacher at some point in their school career. An unfortunate number may have a fairly steady diet of such encounters.

Whether teaching, because of its being relatively protected from public scrutiny, is any more vulnerable to motivational deficiencies (if that is what to call them) among its practitioners than is true of other fields of endeavor is a question worthy of investigation. But even if it turns out that teaching is not unique in this respect, the problem of what to do about it remains. For the influence of such teachers (and again we are referring here to subtle effects) on both their students and their colleagues is almost inevitably detrimental. Teachers who care little about their work or who are not fully convinced of its importance are bound to communicate their feeling to those with whom they have contact. The contagion of such a perspective is well documented in studies of social and institutional decline.

The characteristic response of the typical administrator to this problem is to schedule a "dynamic" speaker just prior to the opening of school or to give a few pep talks of his own, scattered throughout the year. If they are not too saccharine, such inspirational messages doubtlessly help a little. But uncoupled with other efforts their effect seems to wear off rather quickly. What those other efforts might be will depend of course on the particular school and teachers in question; but they might include: altering teaching assignments to create new challenges for teachers who have grown tired of their old way of doing things, introducing inservice programs focussed on professional growth, finding ways of recognizing and rewarding evidence of commitment, increased supervisory support, more visits to classrooms of colleagues, and so on. Vacations have also been known to help!

Closely related to the problem of professional commitment is that of encouraging teachers to become critical of their own actions. Here the needs are both technical and conceptual. We must help teachers see themselves and others at work—through videotapings, observational schedules, more frequent visits to neighboring classrooms—but we must also help provide a critical perspective from which to examine the process of teaching. The latter requirement calls for more than technical know-how. It is closely tied to what educators call "a philosophy of education," by which they mean a way of thinking about what is important in education and, hence, what good teaching looks like.

Judging from all that has been written about teaching, it is probably safe to say that the search for a single critical outlook—a universal view of the good—is foolhardy. Rather, what is required is a wide variety of points of view, similar in many ways to the plethora of aesthetic theories that allow us to examine the same work of art from many different perspectives. In fact, we could probably do much worse than turn to the field of aesthetics for models of how best to proceed. In learning how to make use of such models the insights of the theoretician and the practitioner could be nicely combined.

Finally, there is the problem of how to share more broadly the practical skills of the experienced teacher. Part of this goal can be accomplished simply by affording teachers the opportunity of visiting each other's classrooms more often or by a more careful reading of trade journals, but it is probable that more could be done if larger numbers of educators put their minds to it. One possibility is that of compiling a series of handbooks or manuals that contained practical suggestions (gathered from veteran teachers) suited to different kinds of teaching situations. Such manuals would contain more information of the how-to-do-it variety than does the average textbook used in methods courses in a teacher-training institution. Like cookbooks, these manuals would probably be used rather mechanically by beginners and more selectively and imaginatively by those with greater experience.

The goal of all of these efforts is, of course, to increase the likelihood that teachers will have a profound impact on their students—that their action will make a difference not just in the amount that is learned but in the students' view of what learning and being a learner is all about. We may never be able to describe that difference with great accuracy or apply meaningful numbers to its magnitude, but we can ill afford to pretend that it cannot, and does not, happen.

Chapter 3

A TOOL-DEVELOPMENT STRATEGY FOR RESEARCH ON TEACHING

N. L. Gage

The effort to reveal or explain teacher effects as they occur under present-day conditions should give way to the more fruitful task of increasing those effects. Such an enterprise, if successful, will not merely answer the question, "How do teachers make a difference?" It will improve the amount and quality of that difference.

The enterprise of improving teacher effects implies a style of research and development different from that embodied in much present-day work. In that work, the research is correlational rather than experimental. Correlational research ascertains relationships between variables measured as they occur under natural, or "untampered with," conditions. Experiments get at the relationship between a manipulated independent variable and a dependent variable—one that reflects an outcome of teaching. As compared with correlational studies, experiments in teaching satisfy scientific interests much more adequately because their results indicate much more unambiguously the operation of *causal* influence. And the results of experiments, if they are positive and not readily attributable to chance (i.e., are statistically significant), immediately lead to ways of controlling and improving teacher effects, not merely predicting or understanding them.

Tool-Based Experimentation

But experimentation is not enough. There has been much experimentation in teaching in the sense thus far described. I have in mind a special kind of experimentation. The independent variables, or treatments, or the new kinds of teacher behavior that experiments employ and test, should have certain advantageous characteristics.

Concreteness

First, the treatments should be embodied in materials and equipment. Such treatments have great advantages in changing the instructional situation (Gagné, 1966). Textbooks, workbooks, instructional films, tests, audiotapes, videotapes, programed textbooks, computer-assisted instructional materials, kits, manuals, models, games, simulators, and other devices for arranging instructional experiences in suitable sequences—these are the vehicles through which good influences on what actually happens in schools can be most dependably exerted. Without such material embodiment, attempts to improve teaching and learning run into all the forces that keep people from acting on good educational advice. The advice tends to be too theoretical, too vague as to its meaning for practice, and insufficiently coercive, in the sense that it does not require the teacher to change his ways. Materials and equipment spell out the advice in practicable terms. If properly designed and accompanied with adequate instructions for use, they well-nigh force the teacher and student to do what is wanted of them by the experimenter.

Psychological Validity

Second, the new independent variables must possess psychological validity, i.e., reflect what we know about the factors and processes in learning. The most influential and perhaps most usable conceptions of learning are those embraced by operant conditioning theory. That theory specifies that the teacher should pay attention above all to the reinforcers in any learning situation. The teacher should provide reinforcers to strengthen desirable kinds of behavior and withhold them to weaken the undesirable. So the independent variables in experimental research on teaching should be developed with due regard to the central significance of reinforcement. They should embody attention to allied concepts, such as discriminative cues, preexisting operant levels, extinction, generalization, discrimination, reinforcement scheduling, and shaping.

Another kind of psychological validity can be derived from the structure of what is to be learned, remembered, or applied. Some arrangements of things to be learned are better than others. These arrangements may be called more "meaningful," better "structured," "mnemonically" aided, or "algorithmically" simplified. They take advantage of innate processes and previous learning so as to facilitate the work of the teacher and the student. The materials and equipment provided the teacher should embody what is known about these ways to improve the teacher's behavior and, through this, the student's learning.

Relevance to New Roles of Teachers

Third, the experimental variables should be relevant to the emergent roles of teachers. The technological revolution still underway in education—especially the salient called programed instruction—is making certain tasks of teachers less important. It is also heightening emphasis on other tasks. Programed instruction takes several forms: programed textbooks, computer-assisted instruction, individually prescribed instruction (IPI), and the "program for learning according to needs" (PLAN), among others. All these varieties diminish the teacher's role in communicating and inculcating knowledge—in the sense of increasing the student's ability to recall or recognize facts, definitions, rules, principles, formulas, and the like. They also take over much of the teacher's job of fostering other intellectual skills, such as ability to translate, analyze, synthesize, evaluate, interpret, extrapolate, and apply.

If programed instruction can do all these things, and do them well, what is left for the human teacher to do? Hilgard (1968) offered a threefold answer: the teacher can (a) serve as a model to be imitated in improving the student's tendency to initiate inquiry on his own; (b) provide the kind of approval that only a human being can provide in helping the student develop a favorable view of himself, his learning ability, and his creativity; and (c) foster the student's effectiveness in dealing with other people, in cooperating, sharing, leading, following, resolving conflicts, and tolerating the frustrations of the social world.

Stukát (1970) concluded from his survey of predictions and actual findings on changes in the teacher's role that teaching will entail increased emphasis in the future on continuous diagnosis and evaluation of individual students, counseling and guiding students on short- and long-range plans, interacting with individual students and small groups, and team work and task differentiation among teachers.

Kersh (1965) holds that certain objectives are most readily attained through human instruction, rather than automatic or self-instruction. Among these are patterns of behavior occurring at unpredictable intervals and reflecting "mediational" processes. Examples would include problem formulation, or restructuring, and hypothesis formation, along with their component behaviors, such as shifting, being flexible, and searching for patterns.

None of this means that all older and well-established functions of teachers are altogether obsolete. In their work with individual students and small groups, in stimulating inquiry, and in serving as a model, teachers will still need to be skilled in listening to and understanding a student's question. They will still need to ask enlightening or provocative questions. They will still need to provide clear, extemporaneous, oral explanations of the processes operating in a phenomenon, so that the student need not always discover everything for himself. Accordingly, the experimental variables in research on teaching will still need to be aimed at improving the teacher's questioning, explaining, and listening skills.

Differentiation by Type of Learning

Fourth, the experimental manipulations of teacher behavior should be appropriate to different types of learning. The categories of learning that should be distinguished have taken several different forms over the years. Psychologists like Lewin and Tolman have offered different categorizations. More recently, Melton (1964) edited a volume based on six categories of human learning (conditioning, rote learning, probability learning, skill learning, concept learning, and problem solving), chosen because they were frequently used by research workers. With greater attention to what goes on in schools, Gagné (1971) identifies five "domains"—motor skills, verbal knowledge, intellectual skills, cognitive strategies, and attitudes—that require different kinds of learning and teaching.

These kinds of learning should be taken into account when designing tools for improving teacher behavior. If repetition is more important in learning motor skills than in acquiring verbal knowledge, teachers ought to behave accordingly. If, as Gagné indicates, having a meaningful context is more important in acquiring verbal knowledge than in intellectual skills, again teachers should be constrained to behave appropriately.

Making Average Teachers Able to Teach Well

The new kind of experimental variable should also reduce the demands that teaching imposes upon practitioners of the art. Teaching is less effective than it ought to be because it requires skills, abilities, habits, and powers possessed by only a small proportion of the hundreds of thousands of adults needed as teachers. Teaching now requires levels of sensitivity, in listening to students, that only clinical psychologists and psychiatrists can routinely supply. It now requires adaptability and intellectual agility, in discussions with students, at levels that only professional debaters, trial lawyers, and parliamentarians can regularly attain. It requires the quick invention of definitions, explanations, and justifications, in classroom discourse, according to rules that only professors of logic can adhere to.

Other professions and crafts give their practitioners whole arrays of techniques, instruments, tools, devices, formulas, strategies, tactics, algorithms, and tricks of the trade. Engineering, medicine, law, and journalism, to name just a few, have all of these kinds of aids that make the job possible for ordinary mortals. The engineer has his slide rule, transit, and handbook of stress tables; the physician, his plethysmograph, sphygmomanometer, and pharmacopoeia; the lawyer, his codes, classified collections of precedents, and interrogation skills; the journalist, his formulas for writing leads and his standard rules of content and style.

But in teaching we find relatively few of these ways of making complex tasks more manageable. Teachers are expected too often to rediscover for themselves the formulas that experienced and ingenious teachers have acquired over the years. Each generation of teachers benefits too little from the inventions of its predecessors. The wisdom of the profession does not get saved and passed along well enough for the benefit of the novice. What teaching needs—if it is to be improved in the hands of ordinary persons, who are not geniuses or inspired artists, and if it is to be improved with resources at a level not inconceivably high—is much more abundant and helpful "tools of the trade."

The term "tools of the trade" has appropriately unpretentious connotations. It suggests not theoretical perspectives but quick and easy guides for asking questions and answering them; not conceptual frameworks, but easily applied rules of behavior and performance; not an emphasis on the complexities, subtleties, and profundities of the teacher's task in understanding and helping his students, but ways of making the task manageable. The tools must be usable by persons with the intellectual and emotional makeup that we can expect to find in two million teachers. What teachers need is a reduced demand for arcane insight and creativity and a greater supply of mundane tools.

Programing: An Inflexible Approach

During the last 15 years, the programing approach has been offered to reduce the complexities and unmanageabilities of teaching in just the ways envisaged here. Programing helps the teacher cope with the problems of individualization and cognitive complexity. Individualization adjusts the rate and manner of instruction to differences among students in ability and personality.

Complexity arises in gearing instruction to the twists and turns of cognitive processes—inductive and deductive reasoning, defining, explaining, justifying, and the like. By permitting students to work at their own speed and by arranging what is furnished the student into meticulously planned sequences of steps, programed instruction goes a long way toward handling individual differences and cognitive complexity.

But no one claims that programed instruction will make human teachers unnecessary. As we have seen, even after programed instruction has gone as far as it can go, human teachers will have important work to do. It is with that work, with making it more manageable and effective, that we now wish to deal.

The programing approach has been extended into the realm of the human teacher's tasks. Lecturing, tutoring, and classroom teaching itself have been programed. Let us look at some of these efforts. We consider them here as attempts to solve the problems of making teaching more manageable by ordinary persons. Before evaluating these approaches, we describe them.

Programed Lecturing

Lecturing is hard to do well, and even when it is well done, everyone finds fault with it. It violates the assumptions that good instruction should provide feedback to the teacher and the student, should be adjusted to individual differences, and should entail activity on the part of the student.

The programed lecture (McCarthy, 1970) was intended to remedy some of these defects. These medical school lectures were based on multiple-choice questions shown one at a time via 35 mm. slides and also on sheets distributed to the students, one question to a page, with space for the student's notes. Ten or more questions were presented, in order of increasing difficulty. The earlier questions dealt with basic principles, and the later ones with applications. Pictorial material was presented as needed, with another slide projector.

The lecturer began by showing the first question and asking the students to answer it on their own sheets. Then he discussed the question fully. The students made notes on their sheets for that part of the lecture. Then the lecturer went on to a new projected question, discussing it fully, explaining the correct response, and discussing each possible incorrect response. The content was carefully arranged, moving toward increasing levels of achievement of the objectives. If the student made a correct response to the question, he received immediate confirmation; if he made an incorrect one, he quickly received corrective feedback and remedial instruction from the lecturer.

Such a programed lecture does solve, in a sense, the problem of cognitive complexity that is so difficult to handle in the give-and-take of classroom discourse. But, as McCarthy recognized, the problem of individualization remains, since "remedial instruction is not provided for each individual, as the questions are discussed by the lecturer speaking to the whole audience." Further, the teacher receives no feedback inasmuch as he has no way of telling how well the students are grasping the ideas. According to McCarthy, the feedback to the student may mitigate this disadvantage. In any case, medical students have reacted favorably to the method. More important, this approach provides an example of a concrete tactic that can enable masses of teachers to lecture with greater effectiveness. The technique simplifies some otherwise awesomely complex aspects of lecturing.

Although McCarthy gives no evidence as to the advantage of the method over ordinary lecturing, some relevant findings have been provided by Berliner (1968). Of his three experimental groups, one received training questions approximately 2.5 minutes apart during the lecture; a second, at 5-minute intervals; and a third, at 15-minute intervals. The total number of questions was the same for all groups. A fourth group took notes during the lecture, and a fifth was merely instructed to pay attention. The lecture was presented by videotape to college freshmen in a psychology course. Although the results were complex, the use of the test-like events every 2.5 minutes of lecture produced substantial improvements in immediate test performance.

Programed Tutoring

Tutoring is the teaching-learning situation in which each teacher has one student. Tutoring is frequently used for supplementing the educational program of students from low-income areas. Students who are not doing as well as they ought to, in the opinion of their parents or their teachers, receive tutoring from their parents, from older students in their own schools, or from students in nearby colleges. Typically, the tutors are persons untrained in teaching except for the training they receive as part of the tutoring program.

Can professional and nonprofessional trainers, using prescribed training procedures, improve the performance of upper-grade elementary student tutors? Harrison (1969) had two professional educators and two nonprofessionals at each of five elementary schools train student tutors (5th- and 6th-graders). The carefully prescribed training procedures were aimed at getting the tutor to put the learner at ease, clarify the prescribed task, teach the child how to verify his answer, have the learner read each problem aloud, have the learner mark his answer before providing any feedback, have the learner verify his answer, avoid punishing behavior, provide the learner with verbal praise when appropriate, reward him when appropriate, and check for mastery on designated problems. The trained tutors worked with 1st-graders on "additive sentence equations." A test given during the week after the tutoring showed that the 1st-graders taught by the trained tutors had learned substantially more than similar children taught by untrained tutors. Thus, tutors equipped with explicit techniques were much more effective.

The most thoroughly programed kind of tutoring has been developed by Ellson and his coworkers (1965, 1968). Their work is aimed at developing a technique useful in teaching elementary reading. Nonprofessional persons are trained to follow operational programs which specify in great detail how the teaching is done, and content programs which specify what is taught and the order in which it is first presented. The tutors are "programed" to emphasize success, reinforce correct responses with suggested appropriate words, ignore failures, and go on without comment to the next procedure – all the while recording the child's responses. To observers, the program is hardly visible; the situation seems similar to that in traditional tutoring. Subsequently, the highly systematic character of the tutor's behavior becomes apparent. "An experimental psychologist might see it as a complex and flexibly modified form of the paired associates anticipation procedure, supplemented by verbal approval as a form of reinforcement" (Ellson, et al., 1968, p. 315). The program has loops, short steps, prompts, and branches. The discovery method is used: The first step presents the problem or task in its most difficult form with a minimum of context, and later steps progressively simplify and provide hints or additional information until the child discovers the solution for himself. Professional teachers, who attempted to learn about programed tutoring through practicing it, felt that the program "did not allow them to teach." Apparently, these teachers wanted to give answers or help more rapidly than the program allowed and lacked the patience that the discovery method imposed on the nonprofessional programed tutors.

Ellson and his coworkers have amassed impressive evidence that programed tutoring permits persons with only a high school education and no other preparation as teachers to provide highly effective supplementation of traditional classroom experience in the 1st- and 2nd-grade reading curriculums. As contrasted with "directed tutoring," which is derived from current teaching practice, programed tutoring produced significant improvement in reading achievement test scores when given twice daily (but not when given once daily).

The significant point in the present context is that programed tutoring exemplifies the "tools of the trade" that can ease and improve the work of the teacher. The teacher of the future may serve as a tutor, working with one or two students at a time while the others are occupied elsewhere. This role, Stukát's study suggests, will be an increasingly important one. If so, highly structured, programed tutoring could assume increasing importance.

Programed Classroom Instruction

Classroom teaching is by far the most prevalent mode of teaching in American elementary and secondary schools. Accordingly, the attempt to provide classroom teachers with tools of their trade is extremely important.

Kersh(1965) has attempted to program classroom teaching. First, as already noted, he distinguishes between objectives that could be attained by students working alone with programed instructional materials and those that could be best achieved with the assistance of a teacher. He also refers to "compounded" objectives, those suitable for automatic or self-instructional techniques alongside others that call for human instruction. To attain such compounded objectives requires, in Kersh's opinion, the capabilities of an automated classroom. "Otherwise the teacher would be taxed beyond his ability in the attempt to control the experiences of the learners. To reduce the burden on the teacher and to allow him to concentrate on those activities which require human guidance, a systematically developed set of instructional materials and validated procedure also must be available" (Kersh, 1965, p. 346).

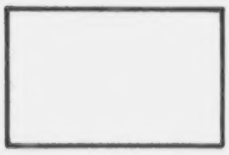
Accordingly, Kersh has developed a notation and charting technique with which the programer could prepare a detailed outline of the learning experience, specifying practice and reinforcement schedules, criteria for branching, and the like. The teacher is thus trained to work alongside the Teaching Research Automated Classroom (TRAC), which provides projectors housed in the students' desk-units and a classroom communication system controlled automatically. Permanent records of each student's performance are made, and class summaries are immediately available to the teacher.

A flow chart, using a special notation developed to indicate specific teaching operations, provides detailed instructions and materials. Figure 1 (taken from Kersh, 1965, p. 354), explains techniques used in the flow chart.

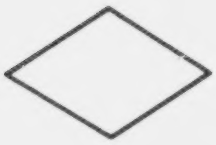
Such a flow chart is illustrated in Figure 2, which specifies the plan for teaching the idea that "A single quantity may be written in different ways."

The first box indicates to the programer that the instruction should start with an example concerning a boy with several

Figure 1.
Examples of Flow-Chart notation
(Kersh, 1965, p. 354)



Instructions to the teacher to explain, identify, or question. Usually may be recorded on tape for playback to children.

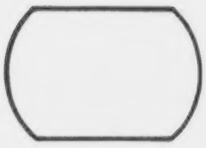


Problems or examples (e.g.) for children. Must be presented to children so they can indicate whether they "solved" it or "not," answer "yes" or "no," etc.

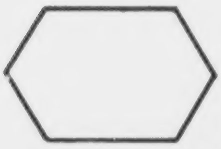
Special notation: ".95+" means 95 percent must get correct answer.

"3 (1.0+)" means continue examples or problems until class achieves 3 in succession, 100 percent correct.

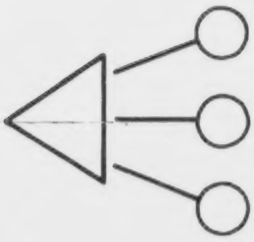
"1.0+" means 100 percent must be correct.



Special instructions for obtaining feedback from children.

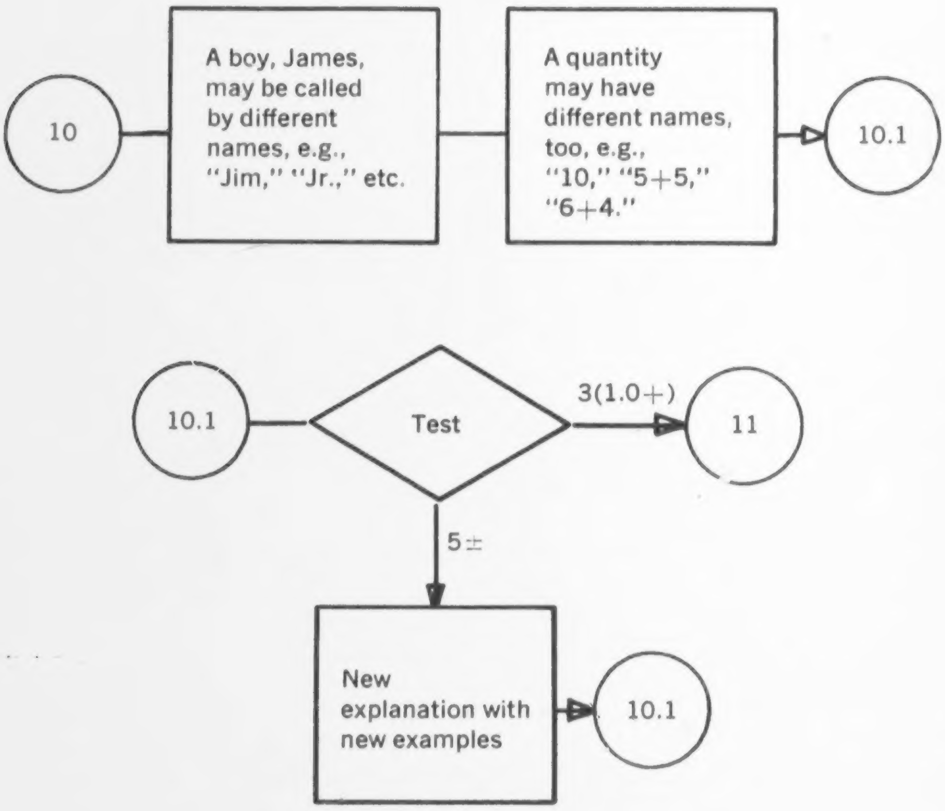


Something the children must "discover" for themselves. Sometimes followed by special procedures to be followed. Otherwise, programer may employ any appropriate technique of nondirective (inductive) teaching.



Branching junction used where sequential order of instructional steps is not important.

Figure 2.
Plan A for Subfact 10,
"A Single Quantity May Be Written Different Ways"
(Kersh, 1965, p. 358)


10
A boy, James, may be called by different names, e.g., "Jim," "Jr.," etc.
A quantity may have different names, too, e.g., "10," "5+5," "6+4."
10.1
10.1
Test
3(1.0+)
11
5±
New explanation with new examples
10.1

> names. The next box indicates that, by analogy, the rule is to be established that a quantity such as 10 may be referred to in a number of different ways including '5 + 5,' and '6 + 4.' At circle 10.1 (simply a location point or connecting link), the flow chart moves to a diamond-shaped frame which indicates that the program should follow with a test. The notation '3(1.0 +)' specifies that the learners should continue with the test until they achieve three examples in succession correctly. The notation '5 ±' specifies that should the learners fail to achieve the criterion after five examples, the program should return them to a new explanation of subfact 10 and then test them again. Circle 11 indicates that when the learners pass the test they are to go to subfact 11 (Kersh, 1965, pp. 355-356).

Kersh spells out the approach in some detail; the present sample merely illustrates it. It represents an attempt to plan teaching carefully so as to reduce the need for creative improvisation.

Needed: An Alternative to Inflexibility and Chaos

The great virtue of programed learning, tutoring, and classroom instruction is specificity. These innovations reduce the amount of chaos in what the teacher does. But such approaches go too far in a good direction. Because they spell out both the content and procedure of the teacher's work in great detail, they impose too much inflexibility on the teacher. Their procedures are topic-specific and must be changed whenever the topic is changed. What is needed are teaching procedures that are general, or usable, over many kinds of topics. Teachers cannot accept complete regimentation through programing of their behavior. What they teach requires more room for spontaneity, creativity, and artistry than such programing allows.

Present-day classroom work, with its "stray thoughts, sudden insights, meandering digressions, irrelevant asides, and other unpredicted events" (Jackson, 1968, p. 4), also makes life hard for teachers. The unpredictability and lack of order become intolerable and eventually lead to an inflexibility of a different kind. The teacher escapes from them to a monotonous acting out, and reenacting, of the same unimaginative and sterile pattern. Thus the evidence, as marshalled by Hoetker and Ahlbrand (1969), indicates that there has been a "remarkable stability of classroom verbal behavior patterns over the last half-century, despite the fact that each successive generation of educational thinkers, no matter how else they differed, has condemned the rapid-fire, question-answer pattern of instruction" (p. 163). As Bellack and his coworkers (1963) describe it, "The core of the teaching sequence found in the classrooms studied is a teacher's question, a pupil's response and, more often than not, a teacher's reaction to that response" (p. 158). Teachers fall into this rut, and stay in it, despite our teacher-education programs, and do what comes naturally, as if they had never been trained. For some writers (e.g., Stephens, 1967) this kind of teaching represents a spontaneous tendency on the part of humans in the role of teacher—a way of behaving that can probably be traced back to ancient times and can be found nowadays on the part of anyone—older child, parent, or professional teacher—who is placed in the role of teacher.

In short, the freedom of nonprogramed teaching turns out to be spurious. Imprisoned by their technical poverty, teachers tend to do the same thing, no matter what they are teaching, no matter whom they are teaching, day after day and year after year.

We need a happy medium between the excessive systematization of topic-specific programed lecturing, programed tutoring, and programed classroom teaching, on the one hand, and the spurious freedom of the opportunistic and unpredictable present-day classroom that Jackson describes. We cannot accept the proposition that the way teaching is is the way it ought to be.

Tools of the Trade

If present-day classroom teaching is too planless and chaotic, and if programed teaching is too inflexible, what is the alternative? The answer here is that research and development should be devoted to the invention, refinement, and widespread distribution of tools for teachers and trainers of teachers. Such tools would vastly enhance the teacher's collection of things to do and the trainer's ways of training him to do them. Such tools would be applicable to many topics, contents, and subject matters.

The improvising musician can create and indulge in flights of artistry only because he has great technical command of his instrument. He can play ruffles and flourishes, crescendoes and diminuendoes, growls and whimpers, pure notes and chords, trills and sustained notes, fast and slow. These skills can be applied to almost any piece of written music and to composing unwritten music on the run, or improvising, as well. Such a musician is not forced—as is the untrained beginner—to pick out the same tune, haltingly and with error, again and again. Teachers need comparable tools so as to be comparable artists. Just as the musician's technical competence frees him for artistry, the teacher's competence in handling his own tools will free him to work artistically in the classroom. Just as improvising musicians can extemporaneously compose variations on a theme, so teachers need tools with which to adapt their behavior extemporaneously to the minute-by-minute variations in classroom topics.

The necessary tools are of two kinds: those for the teacher himself, and those for the trainer of teachers. They can be embodied in materials and equipment. In some cases, the teacher will eventually abandon the concrete versions, just as the child learns to get along without the training wheels on his bicycle. Then the tools take the form of the skills, models, and rules that remain with the teacher after the materials have done their job. Let us now look at some of these.

Tools for Teachers and Trainers

Tools for teachers can take such forms as technical skills, decisionmaking skills, and various kinds of rules, models, and aids. A strong beginning has been made in developing such tools. Nearly a decade of work, initiated by Frederick McDonald, Dwight Allen, and Robert Bush, has resulted in the formulation and definition of an array of technical skills in teaching. Berliner (1969) has provided an account of much of this work. Some of these skills—higher order questioning, reinforcement, probing, varying the stimulus situation, providing silence and nonverbal communication, and skills for controlling small-group discussions—have been fairly well defined in research on teacher training. In that research, different kinds of treatment for trainees have been evaluated through experiments, and their effects on subsequent performance in the laboratory and the classroom have been measured. Those experiments have given us assurance that technical skills can be defined and acquired by teachers. The evidence of their effect on students is still inadequate. But the lines of further work on these questions are clear. Work of that kind will give us a set of technical skills of teaching whose effects on students, when the skills are used judiciously, are known to be desirable.

As Bush (1965) has argued, beyond technical skills the teacher needs decisionmaking capabilities that will enable him to integrate the skills into desirable teaching strategies and plans. That is, he must not only be able to ask higher order questions, for example, but he must also know when it is desirable to do so. He must not only be able to reinforce participation, but also know when it is better for him to hold the stage.

The teacher also needs aids, models, and rules. In diagnosing and evaluating the work of individual students, he can use aids in the form of tests and other diagnostic tools that are easily administered yet usefully valid. In prescribing the student's next steps, he can use charts, tables, guides, and checklists that embody what has been learned from research about the level of difficulty, variety, precision, or discovery that the student's next steps should incorporate. Enormously complex procedures have been made manageable for aircraft pilots and surgeons by means of checklists of all kinds. In the same way, the intricacies of making decisions about what should be prescribed for students can be simplified for ordinary teachers through the use of such devices.

The teacher whose class is engaged in small-group discussions can be guided by flow charts that portray the typical or ideal flow of an argument or discussion on a given topic. He can be aided by forms that help him keep track of the frequency of participation by students. He can be assisted by postmeeting reaction sheets on which students can communicate their evaluations of the discussion in which they have just participated. His work can be made more manageable by the use of algorithms that facilitate his own or his students' analysis of the logic or thoroughness of a groups's problem-solving effort.

In teaching by telling, as in lecturing or explaining, the teacher can be helped by checklists that will remind him of research-based rules (such as a rule-example-rule pattern of discourse). He can use simple devices for obtaining instantaneous feedback that reflects his audience's comprehension, such as cards to be used by students for indicating answers to questions posed by the lecturer or for indicating a need for repetition or further elucidation. The teacher can also be helped with outlines that remind him to organize his presentation along sufficiently redundant and logical lines.

An Example of Tool-Oriented Research

Tools of the trade will have value not merely as the outcomes of research and development. Rather, they should be the entry point for such work. The independent variables in the experimental work on improving the effects of teachers on their students are themselves the tools—technical skills, decisionmaking skills, diagnostic devices, behavior guides, checklists, rules for guiding behavior, and devices of all sorts—that will emerge from such work.

An example of such work will clarify what is being advocated. After several investigations in which my coworkers and I sought behavioral correlates of effectiveness in lecturing or explaining, we turned to an experimental study. In this experiment, we sought a method for improving explanations rather than merely knowledge of the correlates of their effectiveness. So, after having read some treatments of the logic of explanation, Robert Miltz (1971) and I developed a manual, *How to Explain.* The manual and instructions for its use in self-training were given to an experimental group of 30 teacher trainees and were withheld from a control group, who engaged in some other useful activity. The manual presented some relatively simple rules for explaining. First, they were given some simple rules for improving their ability to listen to a student's request for an explanation. Second, the trainees were told they should look for the "things," or elements, involved in the process or phenomenon to be explained. Third, they should identify the relationship between those things. Fourth, they should show how that relationship was an instance of a more general relationship. Finally, they were given some simple rules for the pattern of their discourse, such as the rule-example-rule pattern that seemed in one previous study to be a correlate of effectiveness in explaining.

The manual was used by the teacher trainees working in two-man teams. Each team had a tape recorder, and after one member of the team had practiced a given step, the other criticized and discussed his performance. Then they exchanged roles. They worked through the manual in five 1-hour sessions on each of 5 school days. The evidence from ratings (see table 1) and content analyses (including some made by Jack Hiller with a computer) of the trainees' pretest and post-test explanations indicated that the manual and its accompanying instructions for use brought about a substantial improvement in the explanations of the experimental group. It also revealed some shortcomings in the manual that can be remedied in its revision.

This experiment yielded more than theory-relevant findings to be reported in a scientific journal. In addition, it led to a tool usable in improving teacher behavior. Whether that improvement will last, whether it will show up in actual classrooms, whether it would

Table 1. Mean ratings of pretest and posttest explanations by experimental and control groups (based on Miltz, 1971)

	Experimental group (N = 30)		Control group (N = 30)	
	Pretest	Post-test	Pretest	Post-test
Organization	2.92	3.62	3.29	3.33
Clarity	2.79	3.59	3.23	3.26
Quality	2.84	3.68	3.31	3.34

NOTE.—Ratings of explanations by 10 judges were made on a 5-point scale (5 = excellent, 4 = good, 3 = average, 2 = poor, 1 = very poor).

Differences between experimental and control group Post-test means, both unadjusted and adjusted by analysis of covariance, were significant at the .01 level.

produce improved student achievement—these are questions for the future. But they can be answered readily by the same experimental approach as that already used. And the outcome will be a set of tools, i.e., validated manuals, procedures, and rules for teachers engaged in giving explanations, that can be widely transported and installed in dependable ways.

The foregoing single example is, of course, not the only one that could be cited. Much tool-oriented research has already been done elsewhere, especially in research and development centers and regional laboratories. It seems fair to say even now, in the early years of this approach, that a research and development program of this kind will produce tools that will significantly enhance the quality and amount of the difference that teachers are able to make.

References

Bellack, A. A., and Davitz, J. R. *The Language of the Classroom: Meanings Communicated in High School Teaching.* New York: Institute of Psychological Research, Teachers College, Columbia University, 1963.

Berliner, D. C. "The Effects of Test-like Events and Note-taking on Learning from Lecture Instruction." Unpublished doctoral dissertation, Stanford University, 1968.

__________. *Microteaching and the Technical Skills Approach to Teacher Training.* Technical Report No. 8. Stanford University Center for Research and Development in Teaching, 1969.

Bush, R. N. "The Science and Art of Educating Teachers." In S. M. Elam (ed.), *Improving Teacher Education in the United States.* Bloomington, Ind.: Phi Delta Kappa, 1966, pp. 35-62.

Do Teachers Make a Difference? A Report on Recent Research on Pupil Achievement. U.S. Department of Health, Education, and Welfare, Office of Education. Washington, D.C.: U.S. Government Printing Office, 1970, OE-58042.

Ellson, D. G., Barber, L., Engle, T. L., and Kampwerth, L. "Programed Tutoring: A Teaching Aid and a Research Tool." *Reading Research Quarterly,* 1965, 1, pp. 77-127.

Ellson, D. G., Harris, P., and Barber, L. "A Field Test of Programed and Directed Tutoring." *Reading Research Quarterly,* 1968, 3, pp. 307-67.

Gagné, R. M. "Domains of Learning." Presidential address presented to the American Educational Research Association, New York, February 7, 1971.

__________. "How Can Centers for Educational Research Influence School Practices?" In *Organizations for Research and Development in Education.* Bloomington, Ind.: Phi Delta Kappa, 1966, pp. 21-31.

Harrison, G. V. "The Effects of Professional and Nonprofessional Trainers Using Prescribed Training Procedures on the Performance of Upper-grade Elementary School Tutors." Unpublished doctoral dissertation, University of California, Los Angeles, 1969.

Hilgard, E. R. "Psychological Heuristics of Teaching." Paper for symposium of the National Academy of Sciences, California Institute for Technology, October 28, 1968 (mimeo).

Hoetker, J., and Ahlbrand, W. P., Jr. "The Persistence of the Recitation." *American Educational Research Journal,* 1969, 6, pp. 145-67.

Jackson, P. W. *The Teacher and the Machine.* Pittsburgh, Pa.: University of Pittsburgh Press, 1968.

Kersh, B. Y. "Programing Classroom Instruction." In R. Glaser (ed.), *Teaching Machines and Programed Learning, II.* Washington, D.C.: Department of Audiovisual Instruction, National Education Association, 1965, pp. 321-68.

McCarthy, W. H. "Improving Large Audience Teaching: The 'Programed' Lecture." *British Journal of Medical Education,* 1970, 4, pp. 29-31.

Melton, A. W. (ed.) *Categories of Human Learning.* New York: Academic Press, 1964.

Miltz, R. J. "Development and Evaluation of a Manual for Improving Teachers' Explanations." Unpublished doctoral dissertation, Stanford University, 1971.

Stephens, J. M. *The Process of Schooling.* New York: Holt, Rinehart & Winston, 1967.

Stukát, K-G. *Teacher Role in Change.* Research Bulletin no. 4. Gothenburg, Sweden: Department of Educational Research, Gothenburg School of Education, October 1970.

Chapter 4

STRUCTURE AND TEACHER PERFORMANCE: A PROLOGUE TO SYSTEMATIC RESEARCH

Dan C. Lortie

As I understand our mission, we are here to look for ways to increase our understanding of teaching. Two assumptions appear to underlie the effort—namely, that better knowledge can help raise the performance of public schools, and that improvement is a matter of some national urgency. We are engaged, then, in a kind of research "policy discussion;" that, at least, is the framework within which I wish to present my position.

The argument is presented and developed in three sections. The first deals with general performance levels in teaching, centering on the problem of low teacher concern with the general development of their common art. The thesis, briefly, is this: the structure of the occupation, particularly processes of recruitment, socialization, and reward allocation, presses teachers toward individualism and conservatism. These tendencies, in turn, inhibit the development of collegial involvement in improving the general capacity of teachers to render more effective service. The second section takes up variations in performance by examining ways in which the organization of careers reduces engagement, satisfaction, and, presumably, effectiveness in particular categories of teachers. The final section presents strategies for research which seem consistent with the nature of structures and their reconstruction.

Some may find this paper cryptic, an assessment which could result from an attempt to condense considerable inquiry into a short space. But the intent is not, I wish to make clear, full exposition and analysis of the social system of public school teaching; it is, rather, to illustrate the relevance of structural features to any useful understanding of how research into teaching might help "make a difference."

The General Level of Performance

Occupations differ in their prevailing standards of performance; contrast, for example, the average performance of brain surgeons with the average performance of clerks in retail stores. They also differ in the press members show toward raising the general level of work done by members of the occupation; for example, society grants greater deference to occupations in which colleagues or competitors advance the state of the art and discipline one another to comply with explicit performance standards.[1] Complex mechanisms can develop to reinforce such standards, such as esteem regulation, career disposition, and the like; those contributing to the performance capacity of the occupation may receive its highest accolades. Where such systems work well, there is a steady, internal pressure toward raising the quality of service offered by the occupation.

It is not difficult today to find sweeping condemnations of the general level of performance among classroom teachers, condemnations whose grounds are rarely, in fact, made explicit or carefully documented. But one can be reasonably certain in asserting that teaching is not among those occupations where members play an active part in raising the general performance level of the field. Teachers are loath to hold each other to any set of explicit expectations for performance; classroom teachers, as a group, play little part in advancing the state of practice within their field. Whatever variations in performance may be within the occupation, we do not find a band of superior performers taking responsibility for upgrading general performance. A teacher may be considered outstanding by peers and public alike, for example, without showing interest in the performance capacity of teaching in general. The role, in short, is currently organized to exclude concern with systematic efforts to better standards of performance; in that respect, it is individualistic and, ultimately, conservative.

I shall argue that this particular outlook is not accidental in teaching and that it results, in part, from the way in which our society has ordered the recruitment, socialization, and rewarding of teachers. This is not to say that other factors play no part; one could readily cite, to name a few, the subordination of teachers to administrators and boards, the ecology of mutual isolation found in most schools, even the preempting of intellectual functions by professors in the disciplines and education. But to connect processes of occupational perpetuation (e.g., recruitment, etc.) to the dominant outlook is useful as an illustration of the major point I wish to make. Inasmuch as raising performance levels in general requires the active participation of classroom teachers, the structures shaping their viewpoint are of direct, practical importance. For if we find that the entry, socialization, and rewarding of teachers undermine attempts to increase their involvement in systematic upgrading, it follows that those patterns are part of the problem and had best be part of the solution.

We begin with the process of "recruitment," a term which may connote deliberate attempts to attract persons to an occupation but which sociologists use to refer to all the major ways in which people move into a given line of work. It includes the grounds for individual choice, viewing these in toto as reflecting recruitment resources possessed by the occupation; it also includes the frequently subtle and complicated social pressures which underlie the allocation of people to some part of the work order. Space and time preclude full examination of all these but we can, at least, mention certain aspects which bear on the reluctance of teachers to work collectively on general performance levels.

Historically, American society has never relied on large rewards of money and prestige to attract persons into classroom teaching. In colonial times, payment scales for teachers of the ordinary variety were at or below those of skilled workers; the teacher's status was one associated with the halo of religious function but subordinate to men of the cloth.[2] The development of the modern school system in the 19th century rested on finding thousands of persons to staff a rapid expansion under constraints imposed by limited finances monitored by local citizens. The result was, of course, that teaching became largely the work of young women whose alternatives were severely restricted; teaching became steadily more feminine over the century stretching from the mid-19th to the mid-20th century.[3] Although equal pay for the two sexes prevails today, few economists would deny that the presence of large numbers of women holds down incomes for classroom teachers. Collective bargaining may have moved average incomes to new highs, but current signs suggest that public and school board resistance is stiffening. The central point remains: teaching has never recruited on the promise of a fortune to be made or a high social position to be attained.

Recruitment in teaching has been particularly problematic since expansion needs have been augmented by high turnover. The decision to rely heavily on women was accompanied by the necessity that they leave upon marriage or childbearing; a large proportion of the men who entered were drawn off into administration or other lines of work. High turnover became endemic. How did those whose decisions governed the occupation respond—that is, how can we characterize the recruitment strategy used in teaching?

One major aspect of that strategy has been reliance upon "eased entry," a system in which State governments subsidized and local school systems facilitated the assumption of teaching duties. The States have supported entry through a system of inexpensive, special training institutions, but prior to the 20th century most teachers needed no more than regular school attendance to qualify for positions in the schools. Gradually, the practice of insisting upon college graduation plus special work in pedagogy spread, and some have made much of the gains in education shown by teachers over the last few decades. But as Corwin points out, the relative position of teachers today over the early part of this century is not much, if at all, improved.[4] Nor is teaching defined today within higher education as particularly demanding or difficult of entry. Much of the formal schooling possessed by teachers now, moreover, has been acquired through another special arrangement—the process of "installment schooling" where credits are acquired after work is begun through part-time and summer study.

Observations about eased entry usually turn up in the context of anxiety about the prestige of teaching. But other kinds of effects are associated with this structural arrangement; principally, people have chosen teaching without significant constraint from standards developed by the colleague group. To note that entry is eased is not, of course, to identify the variety of "motives" of those who select the occupation; given the massiveness of teaching, these are undoubtedly diverse. But eased entry does mean that a considerable amount of "self-selection" takes place in the recruitment into teaching; teachers-to-be are not closely screened through a strong, clear image of the qualities needed to perform the role. There is, of course, some screening, but the "net," so to speak, is coarsely rather than finely woven. This characteristic of recruitment is one of the important bases for the "individualism" we find in teaching; there is considerable "play" in the process and relatively little narrowing down to persons of a particular kind.[5]

The special characteristics of teaching, however, result in certain interesting features in the flow of entrants. Teachers are unusually visible to practically all young Americans, thus standing ready to serve as models for identification. That this influences recruitment is supported by the relatively large proportion of classroom teachers, particularly women, who report early and firm decisions to enter the occupation. Elementary teachers in particular are likely to report that "I cannot remember when I didn't intend to become a teacher" or to locate their decisions during elementary and early secondary school.[6] If we array occupational decisions on a continuum ranging from affective at one end to coolly rational at the other, the testimony of these teachers places them at the emotional end; they do not describe their choices as the careful weighing of relative costs and benefits between balanced alternatives but as affected by particular models they encountered. For many women teaching today, the selection of teaching is somewhat like a "calling."

Teaching has another characteristic, however, which draws other people in on very different bases—it is permeable late in one's schooling and can provide a "fallback" occupation when other plans fail to work out. Thus we find that a significant proportion of teachers, particularly men, choose late and describe their entry less as choice than as necessity.[7] Like other men in teaching, these persons aim, in the main, to occupy other than classroom positions.[8] Since some move into administration, it is likely that their viewpoints have influence which exceed their numbers.

One is immediately struck by the tensions that must result from the joint participation of these two particular groups in the same occupational role. But there is irony in that the presence of two such evidently dissimilar subgroups may have similar effects on the subculture of teaching. The affect-laden nature of girlish choices (the predominance, that is, of identificatory elements) points to orientations more individualistic than collegial, more conservative than change oriented, more intuitive than analytic. For each young woman brings with her private models of desirable teacher behavior, a kind of looking to the past rather than present practice or an ideal future; such teachers are likely to experience their teaching acts as "living out" past identifications rather than adhering to canons of a dynamic craft.

The late-deciding men, on the other hand, have a limited stake in teaching per se—they see it as an escalator carrying them to other destinations. This aspiration, naturally enough, reduces their investment in the refined development of teaching practice and in energy-demanding efforts to upgrade the standing of the teacher. Their interests, in fact, may run in a somewhat contrary direction, for they have a vested interest in the continuation of mobility opportunities and may even associate those with continued subordination of teachers. Hierarchical upward mobility involves detaching self from close lateral ties, a process which evokes its particular type of individualization; it is also associated with conservatism on matters of structure and authority.

Thus persons of divergent experience take stances of individualism and conservatism toward teaching on quite different bases. If our analysis is accurate, important subgroups of teachers have little internal press toward collective concern with developing the state of their art. Some of this low interest, it seems likely, stems from the types of recruitment which prevail in American public schooling.

Socialization into teaching has received more attention than recruitment, and here we can begin to rely on studies done by several researchers—studies which are moving toward increasing consensus on the special characteristics of this process in public education.[9] It seems clear that the ways in which people are shaped into teachers articulate with patterns of recruitment; underscore individualism, conservatism, and low commitment to analysis of teaching work; and retard concern with its cumulative development.

A recruitment pattern of eased entry is, of course, the antithesis of strong emphasis on personal changes induced through training. How could thousands of would-be teachers be attracted if they were forced to spend long years preparing for a low-income, short-term career? (One sees, in fact, the persistence of this point of view today in subtle ways; much recent government policy seems directed more toward facilitation than toward intensification of professional preparation.) Nor has American society invested heavily in institutions of teacher preparation, either preservice or during service; teachers have been trained in relatively inexpensive programs and school systems have been very reluctant to spend resources in anything but superficial kinds of inservice training. The result is that observers of teaching (in a comparative context) characterize formal socialization attempts as relatively weak in their impact on students.[10]

The apparently low psychological impact of teacher training means that prior and subsequent influences play an important role in forming teacher attitudes toward their occupation. Although our knowledge is not precise on this matter, there are indications that years of exposure to teachers prior to teaching leave an important imprint. This learning seems to be largely of the slow, unwitting type, a kind of involuntary imitation, often triggered by events which occur later in the teacher's classroom. Again we note a preservation of the past accompanied by the sway of "intuition" as against conscious deliberation in choosing teaching tactics and strategies; to the extent that teacher training fails to disturb earlier learning patterns, it perpetuates what was absorbed as a student.

We are just beginning to study the critical beginning teaching years; studies are appearing now which reveal the play of influences coming from colleagues and from students—it is strange how long it has taken researchers to note the critical role students play in influencing the beginning teacher.[11] Yet it is a matter of simple observation that the beginning teacher, unlike the neophyte in many other occupations, does a remarkable amount of learning outside the presence of other adults and away from personifications of possible criticism and review. Beginning teachers react to stressful exigencies, consciously and unconsciously testing alternative approaches, hacking out a style under circumstances ill-adapted to careful reflection and extensive experimentation and review. The expectations held toward beginning teachers, though probably relaxed in some respects, are formally similar to those extended to experienced, "master" teachers; they are under pressure to prove their capacity to organize instruction, keep order, and do all the other tasks of the teacher. The data gathered by the respondent suggest strongly that teachers emerge from this difficult period with little confidence in the existence of "principles of pedagogy"; what they have learned in becoming teachers, they say, are effective devices, "bags of tricks" peculiarly suited to their own personalities.[12] The individualistic tone is unmistakable—to teach well is to work with one's peculiar self, not to share in a generally valid body of practical and scientific knowledge. It links with the conservatism of remembered models to reinforce the outcomes of recruitment based on "eased entry."

I will take but a few sentences to indicate how the reward system in teaching does little to offset individualism; in fact, it probably amplifies this proclivity. Extrinsic and ancillary rewards (money, rank, security, and favorable schedules) tend toward automatism in teaching; they are, at least, regulated by longevity and coursetaking rather than demonstrated effectiveness in the classroom. I have argued elsewhere, with supporting data, that teachers seek to maximize total rewards by exerting efforts where they can pay off in increases—in the psychic domain of rewards occasioned by the sense of effective transaction with students.[13] Such rewards are experienced in classrooms by individual teachers and are only peripherally influenced by administrators and colleagues; their realization rests less on relations with other adults than on relations with students. Thus the tendency to move toward students and away from other adults in the school setting is encouraged. A structural correlate supports this "privatization": teachers encounter very few occasions for honor from their employers, colleagues, and others generally; there is little or no payoff, in fact, for activities directed toward enriching the common culture of the occupation through diffusing innovations or making fresh observations and the like. Seeking rewards through student response, however, may occasionally offset conservatism—this we would expect, for example, with today's high school students making demands for "relevance" or "nowness" from their teachers; the author's data do not deal with this (although elementary students seem rather conservatizing) and it should be a promising area for research. (It may be that change takes place through a cycle of influencing students first, teachers later.) But if students press teachers toward change entirely on *their* terms, it is not at all certain that such effects will increase the instructional potency of teachers; in fact, we may simply be witnessing more teaching of teachers by students, a process we have undoubtedly underestimated and one whose effects are not knowable a priori. The key point remains: The current distribution of rewards in American public schools is not geared toward increasing the sensitivity of teachers to their common problems and the generation of shared solutions.

Much more could be said about the source of teacher resistance to collegial mutual involvement, shared concern for enhancing technique, and the apparent low interest in creating a dynamic craft. But it should now be clear that much more is involved than sheer cussedness or laziness—the themes in so many ill-informed attacks on teachers and schools. The sentiments of teachers are themselves produced by American society; the society develops particular forms of attraction, induction, and reward for those in classroom teaching. Those processes, it appears, lead to individualism and conservatism—to the opposite, in short, of eagerness for change and deep commitment to the occupational weal. I hope this argument is sufficiently persuasive to commend further research to reject or amplify it, for if it holds truth, efforts to raise general levels of performance will have to come to grips with such processes. But any such steps should not be taken blindly; we need reliable and detailed knowledge to stand an even chance of successful intervention.

Variations in Performance

It is probably easier to convince people that the structure around an occupation influences its general nature than it is to demonstrate that variations in performance also have structural antecedents. Persons trained in science will quickly sense the contradiction of suggesting that a constant produces variation and they are, of course, right. But structures can influence variation when one adds additional variables which reveal the interplay of structures and human lives; I shall try to sketch out such interconnections with illustrations from research. The purpose is again heuristic, for these are complex questions requiring considerably more inquiry; my hope is that the analysis, despite its necessary tentativeness, is sufficiently persuasive to show that further research is desirable.

I begin with a commonplace observation—there is a near lack of any significant career progression in the work lives of those who stay in classroom teaching. "Near lack" is stated because the career possesses some slope; there are the incremental, annual gains; the possibility of improvement through mobility (a not-too-relevant consideration, of course, for the modal teacher who is a married woman); the chance for a department chairmanship which, in some systems, marks genuine status change. In the main, however, the teacher who has attained tenure rank is unlikely to experience significant career steps after that point. Compared to career systems in practically every other kind of organization, the early and late status of the person who stays in teaching are remarkably similar.

It appears that this characteristic of the teacher's work life has implications for variations in performance found within the occupation. Beginning with men who teach, we have already noted that most beginning male teachers aspire to work which will take them out of the classroom. Many will undoubtedly achieve their goals—higher ranking positions within education are disproportionately occupied by men. But no mobility system has, to this writer's knowledge, ever worked to satisfy the hopes of all; for whatever reasons, many will not manage to attain higher position within education or careers outside teaching. From interviews conducted with older teachers who have not made it out of the classroom, it seems that school people, however, expect men to "succeed" by leaving the classroom; to remain bears marks of stigma, of less than full achievement.[14] Despite all the talk we have heard about the desirability of more men in public school classrooms, it seems that those engaged in school work (and probably the society at large) do not consider teaching an honorable terminal position for a man.

Men over 40 in a sample gathered by the writer typically display feelings which are consistent with such a definition of their fates. They rarely express high satisfaction with the way things have worked out. Some go out of their way to comment on the "dishonesty" and "politics" (in this context, a pejorative term) controlling promotions in their school system. Others show varying degrees and types of withdrawal from their work, ranging from some whose "second jobs" now seem to be their primary commitments to those whose avocations are discussed with greater enthusiasm than marks their vocational talk. In another sample, the persistence of administrative ambition was notably longer among men teaching in elementary schools, a finding which comports with the allegation that teaching young children is even more status incongruent for an older male.[15]

The structure of American schools—in this case, the emphasis on hierarchy and the separation of administration from instruction—puts men who do not ascend or leave in the position of being "passed over," an abrasion exacerbated by the preference shown, with the years, for younger, less experienced men. One cannot but hypothesize that for at least some of these men, psychological effects accrue, effects which are likely to dampen the qualities usually associated with effective teaching. Take, for example, an oft-cited quality, that of "enthusiasm"; can one sustain this precious quality in light of growing fears that one has, after all, failed? How readily can a man whose self-esteem has been shaken convince students that their common activity—learning—is truly worthwhile? By institutionalizing the expectation that older men will not teach (we surely need more information on the extent and scope of that definition), have we not made it likely that an important subgroup of teachers will perform at less than peak effectiveness? I pose this in question form deliberately—we clearly need sensitive, carefully conducted research on this delicate issue.

Persistence in teaching is also complex for women, but here the complications touch on matters of marriage and family. In "the bad old days," schools did not hire married women or retain single ones who wed; the experienced teacher was invariably a spinster. The unmarried teacher may not have been widely envied, perhaps, but she did have a special kind of status within education—she became the "Miss Dove," the symbol of long, dedicated service to children. But today the situation has changed radically for women who stay in teaching and do not marry. The modal teacher is now a married woman; this means that in most faculties, single women constitute a minority, and, increasingly, a group which holds no monopoly on experience. They are, furthermore, made constantly aware of the ambiguities of their social status.

Married women who teach must balance the claims of work with the claims of family responsibilities. Although teaching presents fewer problems of this kind than many alternative occupations for women, we see the balancing effects in the lower participation of married women in extra-class activities.[16] Yet married women, despite their somewhat lower contributions, receive greater rewards; at least they report, on the average, higher rates of satisfaction with teaching than do single women.[17] We have here what is probably a classic instance of both relative deprivation and imbalance in contributions and rewards; single women teachers lose out.

To older men who teach, therefore, we can add another subgroup liable to feelings of less than unbounded joy, to at least occasional propensities toward alienation from work. One must be extremely careful, of course, not to allege that all single, older women teachers fail to get satisfaction from their work—that is simply not consistent with any evidence I know. But there are indications that the lot of some single teachers can be sufficiently difficult to affect their teaching. Among those interviewed for the author, for example, it was found that this group produced more than its share of bitter respondents eager to dissociate themselves from teachers as a group and more older teachers who reported severe and chronic fatigue. Is it not likely that such problems influence daily performance, at least at times? Again, this is a question of some delicacy but nonetheless one of some relevance to any searching examination of variations in teacher performance.

Any connections between structural interactions, personal characteristics, and teaching effectiveness obviously rest upon an important social psychological assumption. That premise is that the attributes we generally believe to be correlated with high teaching effectiveness are liable to situational variations stemming from life circumstances and, chances are, other influences such as organizational climate, community attitudes, and the like. We must guard against a natural tendency, I believe, to think of such attributes as stable, personal properties, as flowing solely from the personality makeup of the teacher. This is not to deny that personality is a matter of consequence—some people's "low points" probably exceed the "high points" of others. But it hardly seems warranted for us to assume, a priori, that the person who begins teaching with enthusiasm, high commitment, and imagination will necessarily so continue or, for that matter, that the awkward beginner may not mature into a high performer. Situations affect personal dispositions; frustrating situations can probably erode the psychological resources of those who tackle what is, after all, a demanding occupation. Our research agenda into levels and variations of performance among teachers, therefore, should include close attention to the phases of lives and their interaction with the particular organization of teaching work.

Research Strategies

It is easy, in "normal times," to dismiss research into basic structures as impractical; one can say that although it is interesting enough, it provides too few leads into feasible, immediate change. Such structures, after all, are tough—they are time-honored, mutually reinforcing, and too diffuse to get hold of. One looks instead for ready mechanisms (e.g., curriculum materials), arguing that knowledge about them is likelier to prove useful in action.

But these are not "normal times" in American education. Many signs point to ours as an era of genuine change, even of the passing of an established moral order. The school systems designed in the 19th century are under broad attack—an attack not only of words but of more and more counteractions. The paternalism of many years ago has dissipated, being replaced by power plays and hard bargaining. Today, for example, teachers, parents, and students handle their relationships with strikes as well as conferences, through political agitation along with respectful petitioning, with lawsuits to replace "dress codes" and other remnants of rejected life styles. Other trends reveal sharp discontinuities with the past, such as contracts to private firms to teach in public schools, new technologies such as computer-assisted instruction, and new conceptions of teacher and student roles, as in "informal education" programs. School buildings, to cite another example, used to look as if they all came from one stencil; today they are designed to bring about new social relationships. It is not merely that doubt is being cast on the legitimacy of the old, standard order; new patterns of behavior are beginning to appear.

Periods of transition are exactly those in which serious analysis of our basic institutions is most relevant to making policy and devising new administrative arrangements. We can see the beginnings of this now, for example, as ideas generated in social science become adaptations in schools; such influence is evident in team teaching or the cadre system under development at the University of Chicago or in the new divisions of teacher tasks being worked out in various school districts. Centers for research and development are establishing continuing relationships with schools. In short, change is coming about both politically and administratively; the question is, what effects are these developments having on teaching and its effectiveness?

The potential dynamism of American schooling, however, is not matched by reliable knowledge on the workings of its social system or on the likely outcomes of various alterations within it. Where events move faster than understanding, it is pretentious to claim that we are planning and shaping that reality. Mindless novelty will be no improvement over the mindless routines castigated by Silberman.[18] Nor can we rely solely on doctoral dissertations, colorful journalism, or occasional scholarly investigations to get our picture of how schools work or how new developments affect them. We need programs of research designed specifically to inform the deliberations of those who will govern events during this key period—citizens, teachers, boards of education, government officials, and others.

Time permits but a few observations on research strategies. The first I should like to make deals with the ideas we bring to bear in such inquiry. The development of deeper understanding of how structures inhibit effectiveness and how other structures might foster it is not likely, I believe, to come from any single scientific discipline. We can not achieve this difficult task if we cling to tribal loyalties and disciplinary myophias. Those supporting such research should seek out contributions from anthropologists as well as economists; historians as well as psychologists; psychiatrists, political scientists, philosophers, and others as well as sociologists. The efforts of such specialists, moreover, should be conjoined with those who have intimate familiarity with school affairs; the study of teaching should include contributions from teachers, counselors, principals, and others who have developed sensitivity to how it operates and how it might operate better.

Secondly, I believe the time for talking about "alternative systems" is over–it is time to institute them. We can, of course, learn much from what exists, but our estimates on the effects of alternative courses of action will be weak until we can test them in action. Innovation processes in public schools are not always, by any means, sufficient for purposes of inquiry; too frequently, school boards are fickle in their attachment to new practices, dropping them before evaluation makes sense. Even where new programs persist, the quality of evaluation is frequently low. To project truly different structures for schools in general, we need stable, carefully designed exemplars; let me illustrate, using notions developed earlier.

Schools today appear to be staffed predominantly by teachers who positively identified with their own teachers (where such identification occurred) and who consequently tend to employ them, often unreflectively, as models for current behavior. Would a school look different if an effort were made to staff it with those who counter-identified and consequently want to depart from their own experiences as students? Would such a staff be more innovative and more ready to engage in steady analysis of its actions? If so, what changes in recruitment policy might bring in more teachers of this orientation? Current training of teachers includes courses where professors seek to instill "principles" for teachers which, it seems, they reject after experience. Could new programs be designed in which students genuinely reviewed their own experience, assessing it carefully and explicitly, and consciously selected desirable and undesirable features from the past while developing greater openness to the new?

Schools currently feature automatic reward systems in which classroom effort and effectiveness are largely ignored and in which contributions to general performance capacity are overlooked. How would teachers behave in a school designed to encourage close analysis of teaching decisions, a school where time, technical facilites, and rewards conjoined toward a focus on learning how better to teach students? Would new and more precise forms of communication among teachers replace the vague, often empty rhetoric so prevalent today? What would occur if we instituted and sustained a pattern of differentiated career as suggested by Benson and others?[19] Could stigma now attached to staying in the classroom be offset by introducing recognition for those who are committed to instructional activities? The questions are easily proliferated; without imaginative and bold action, however, it is unlikely that we will get the answers.

My third and final suggestion is already implicit in what has been said. The kind of research suggestions made presume longer time commitments than are currently fashionable among government agencies and foundations. The problem is, of course, that altering structural arrangements takes time; the careers of persons unfold over years, and testing relationships between different entry, socialization, and reward patterns requires longitudinal studies to assess the effects. School systems are not in a position to undertake and sustain long-term programs of this kind, and the complex causation that lies behind that problem suggests that they will not be able to do so for many years. I would urge, therefore, that government agencies concerned with education amplify the tendency seen, for example, in longer periods of support for research and development centers, and commit resources to research schedules consistent with the nature of teaching careers and organizational reorganization. Education needs no more corpses of immature ideas strewn along its path; it is time to overcome the proclivity toward short-run fads and fashions decried years ago by Sumner and by many other observers since.[20]

To argue that the quality of teaching in our schools is influenced by social arrangements need not be a counsel of despair. For it grants, after all, the possibility that other, superior social patterns will raise our goals and hopes while releasing new potentialities. To me, that is what education is about.

Footnotes

[1] Collegial standards are most relevant, of course, in fully developed professions. But competition can perform similar disciplining functions where performance standards are explicit; this is illustrated by professional athletics and certain types of competitive business situations.

[2] Willard Elsbree, *The American Teacher* (New York: American Book Company, 1939).

[3] David Tyack, *Turning Points in American Educational History* (Waltham, Mass.: Blaisdell Publishing Company, 1967).

[4] Ronald G. Corwin, *A Sociology of Education* (New York: Appleton-Century-Crofts, 1965).

[5] J. W. Getzels and P. W. Jackson, "The Teacher's Personality and Characteristics," in N. L. Gage, *Handbook of Research on Teaching* (Chicago: Rand McNally, 1963).

[6] These observations are based on 94 personal interviews conducted in five school systems in New England. Larger samples gathered by Benjamin Wright support the general points.

[7] New England data. A sample collected by Carol Kronus supports the tendency of men to decide later and after other alternatives failed to work out.

[8] Ward S. Mason, *The Beginning Teacher* (Washington, D.C.: U.S. Government Printing Office, 1961).

[9] Other articles bearing on these themes by the author include "Shared Ordeal and Induction to Work," in Becker, Geer, Riesman and Weiss, *Institutions and the Person* (Chicago: Aldine, 1968); and "Teacher Socialization: The Robinson Crusoe Model," in National Education Association, *The Real World of the Beginning Teacher* (Washington, D.C.: NCTEPS, 1967). The first article cited includes references to several studies of socialization done by others.

[10] Robert Dreeben, *The Nature of Teaching* (Glenview, Ill.: Scott, Foresman and Company, 1970).

[11] Emil Haller, "On Moving to Smaller Rooms," *Administrator's Notebook,* XV, No. 6 (February 1966).

[12] The data and analysis alluded to here will be presented in a forthcoming monograph by the author, tentatively titled "Structure and Sentiment in Classroom Teaching."

[13] Dan C. Lortie, "The Balance of Control and Autonomy in Elementary School Teaching," in Amitai Etzioni (ed.), *The Semi-Professions and Their Organization* (New York: The Free Press, 1969).

[14] New England interviews.

[15] Questionnaire data (N = around 5,000) gathered in Dade County, Fla.

[16] New England interviews and NEA data.

[17] New England interviews and NEA data.

[18] Charles Silberman, *Crisis in the Classroom* (New York: Random House, 1970).

[19] Charles Benson, *The Economics of Public Education* (Boston: Houghton Mifflin, 1961).

[20] William G. Sumner, *Folkways* (Boston: Ginn and Company, 1906).

Chapter 5

NEW DIRECTIONS FOR RESEARCH ON TEACHING

Barak Rosenshine

Our habit in the field of education has been to introduce innovations and then to discard them and introduce others. Although many people are calling for change in American education, change itself is not sufficient. We need to distinguish between alternatives which are merely changes and alternatives which are improvements in the process of education in classrooms. Specifically, we need to identify the features of successful programs which appear to be causing the success. Without such study, education will continue to be susceptible to gimmicks and fads which are offered as innovations.

The focus in this paper is on suggesting new directions for research on teaching. Particular emphasis is given to the need for research on teaching while the teachers and students are using curriculum materials packages which have been developed and are being developed in national centers.

The three parts of this paper—(a) lack of research on how teachers make a difference, (b) tools for future educational research, and (c) suggestions for payoff research in real classrooms—were written with different but overlapping audiences in mind. The first part was written for the general audience of people interested in education. The second was written for the general audience and also for researchers, readers of research, supervisors of research, and social critics. The third was written primarily for educational researchers and supervisors of research, although all readers are invited to read and criticize the material; the general reader will probably find the section titled "Curriculum-Specific Studies: Examples of Research Within Curriculum Implementation" the most interesting, and may wish to limit his reading of the third part to that section.

The third part (c) is redundant in that the same basic designs are presented within four contexts: curriculum-specific studies with and without comparison groups, naturalistic studies, and teacher training programs.

Lack of Research on How Teachers Make a Difference

In both an absolute and comparative sense there is a notable lack of classroom research on how teachers make a difference. That is, there is a lack of research in which an attempt has been made to relate the instructional activities of real teachers and real students to measures of student growth. At most, there are 70 correlational or experimental studies in which observed behaviors of teachers or students have been related to student growth. Almost all of these studies were reported in 1966 or thereafter; approximately half were conducted by doctoral students who had limited resources and so had to use 15 teachers or fewer in their samples. The number of instructional behaviors which have been studied is limited, and many of the activities which are of interest to educators and the public have not been studied to any large extent *in situ.*

The reader is encouraged to check my assertion on the paucity of classroom research on *how* teachers make a difference. For example, he might inspect any recent textbook on educational psychology or subject area instruction. I contend that in both types of book very few of these 70 studies appear. Instead, the majority of references about instruction are to laboratory studies in which small aspects of learning were studied, or to classroom-based studies involving one or two teachers, one of whom was the experimenter. The authors of these textbooks attempt to offer advice to teachers, but the research base is best suited for suggesting future research, rather than mass practice. That is, laboratory studies of classroom-based studies involving a small number of teachers provide an excellent source of ideas for future studies involving a larger number of teachers and classrooms; the results of these future studies will be better suited to provide advice to teachers.

Some of these textbooks do contain reports of studies in which one group of teachers used a special method (such as inquiry method, activity method, or individualized attention), and a comparative group apparently did not use the method. In most studies such as these we have little or no information on the instructional activities which occurred within each method. Unless we know the instructional activities which occurred in the classrooms, we have little information on how using a special method made any difference.

Furthermore, one could verify the paucity of research by reading the abstracts of the papers presented at any meeting of the American Educational Research Association. Of the 1,000 papers presented at the 1971 meeting, no more than 15, by the most generous criteria, can be labeled research on how teachers make a difference. That is, no more than 15 studies focused on how real teachers make a difference on any criterion of student gains. In the 7 years that the American Educational Research Journal has been published, no more than 10 studies have appeared which meet these criteria. Such statements as these seem extreme; the reader is invited to inspect the evidence himself.

Statements on the paucity of classroom research seem strange: (1) after 50 years of laboratory studies and small, experimental classroom studies on research in the teaching-learning process; (2) after a decade of massive investment by governmental and private agencies in the development of new educational programs, particularly when funds for evaluation accompanied many of the programs, and when evaluation is a continual part of programs such as those funded under title I of the Elementary and Secondary Education Act of 1965; (3) following the development of a large number of curriculum-materials packages, after the large-scale research study of 1st-grade reading (Bond and Dykstra, 1967), and concurrent with the study of different programs within Project Head Start and Project Follow Through; and (4) when one considers the millions of dollars spent for preservice and inservice institutes designed to train teachers to use these new curriculum materials packages, or the fact that at least 50 studies appear every year in which the behavior of teachers is described using observational instruments.

It seems to me that the problem has been one of excessive compartmentalization and fragmentation within the field of education. For example, laboratory studies only generate additional laboratory studies, and the results of these studies are seldom applied to classroom practice. Teaching methods are seldom derived from prior research, and new ideas are usually implemented full scale as classroom practices to be followed, instead of first being tested and modified in a research setting. The research conducted on new teaching methods is usually for the purpose of proving an idea, and is seldom conducted for the purpose of improving a method. Hundreds of classroom observational systems have been developed, and numerous studies of current teacher instructional activities have been reported using these systems, but no more than 10 of these observational systems have been validated against any measure of student growth. Curriculum materials packages are developed and disseminated, but there is little monitoring of the instructional activities which occur while the packages are used in real classrooms. Teacher performance criteria are developed, usually quite apart from the curriculum packages, and teachers are trained in these practices, but the effect of such training upon student growth is seldom measured. Finally, sociological studies on teachers, teaching, administrators, and the introduction of new developments are seldom integrated or applied to all of the above problems.

There are several possible causes for the lack of research on how teachers make a difference, two of which will be discussed. The first comment focuses upon the research finding that a curriculum materials package is not a single variable, and the comparable finding that an instructional approach is not a single variable. The second comment concerns the lack of and need for a systematic plan to include research within all phases of development, dissemination, and evaluation.

An instructional program is not a single variable. We are still thinking of instructional programs as if they were single variables. That is, we speak as though the hundreds of programs labeled Head Start, Montessori, or Follow Through are indeed the same program. We speak as though the thousands of classrooms using the same curriculum materials package, such as Individually Prescribed Instruction (IPI), the Biological Sciences Curriculum Study (BSCS) program, or the University of Illinois Committee on School Mathematics (UICSM) program, are indeed receiving the same instruction.

The fact that an instructional program is not a single variable can be demonstrated in a number of ways. One approach is exemplified by the Office of Education's cooperative research program in 1st-grade reading instruction (see Bond and Dykstra, 1967), in which one of the purposes was to determine which of many approaches to initial reading instruction produces superior reading and spelling achievement at the end of the 1st grade. When the results from 27 projects involving 137 classrooms were analyzed, there were few differences among the methods on the various measures of reading and spelling achievement. This lack of significant differences led the investigators to conclude that there are variables other than method which influence student success in reading. These investigators and other reading specialists have suggested that future research should focus upon the teacher, the instructional activities, and characteristics of the learning situation rather than upon method and materials. The number of subsequently published studies on how the teacher, using the methods and materials, makes a difference has been small.

Another way to demonstrate that an instructional program is not a single variable is to focus upon those studies which have investigated the instructional activities of teachers within a single program, or compared the instructional activities of teachers across two or more programs. Such studies are reviewed elsewhere (Rosenshine, 1970), and the reviewer concluded that there are significant differences among the instructional activities of teachers within specific instructional programs or curriculum materials packages. The most dramatic of these studies is one in which counts were made of various types of cognitive activities which occurred in the classrooms of six teachers who were teaching the same unit from the same BSCS curriculum materials package (Gallagher, 1966, 1970). On almost all measures of teacher behavior there were significant differences among the six teachers. For example, in this inquiry-type BSCS curriculum, the amount of teacher talk across the six teachers ranged from 66 to 95 percent. When ideas were being evaluated, teacher talk ranged from 57 to 100 percent; when ideas were being expanded, from 67 to 100 percent; and when ideas were being explained, from 59 to 100 percent. When the percent of "topics" in various dimensions of Gallagher's topic classification system was studied across the six teachers (regardless of whether the teachers or the students were talking), there were equally wide differences. For example, topics on the "data level" ranged from 3 to 32 percent; on the concept level, from 58 to 95 percent; and on the generalization level, from 2 to 16 percent across the six teachers. The investigator concluded:

> The data would suggest that there really is no such thing as a BSCS curriculum presentation in the schools . . . each teacher filters the materials through his own perceptions, and to say that a student has been through the BSCS curriculum probably does not give as much specific information as the curriculum innovators might have hoped (Gallagher, 1966, p. 33).

We are now beginning to recognize that simply developing a curriculum materials package, an instructional method, or an educational innovation is not sufficient; we are now beginning to recognize that studying the way an educational product is used in the schools is at least as important as developing the product. But we have spent too little time and money studying how products are used and modifying products on the basis of such study.

There is a need to join instructional research with program development and dissemination. Perhaps a second reason for the lack of research on how teachers make a difference in real classrooms, with real children using real materials, has been the failure to fund instructional research within the development and dissemination phases of curriculum materials or teacher training programs.

Instructional research within national curriculum programs is lacking. There are several national curriculum programs which are 5 years old or older, which still maintain central offices, and for which teacher training workshops still exist. These include BSCS, UICSM, the School Mathematics Study Group (SMSG), the Physical Science Study Committee (PSSC), IPI, and the American Association for the Advancement of Science program on Science—A Process Approach. Although their curriculum materials packages used by these programs have been disseminated in thousands of schools, and although teacher training workshops have been conducted for almost all of these programs, only one central office, to my knowledge, has developed an "official" observational instrument or student questionnaire which can be used to determine whether the instructional activities are in accord with the intentions of the program developers. The single exception, the IPI program, has developed such an instrument, but according to the most recent report I have seen (Lindvall and Cox, 1970), its use has been limited to four classrooms.

In the main, there has been little research within these national curriculum programs on the instructional activities which occur once the developed materials have been disseminated, and even less research on the relationship of these activities to expected or unexpected outcomes. Little time and money have been devoted to learning what occurs as a curriculum package is being used, and few data have been gathered which can lead to revision of the program or the original philosophy.

In addition, I have found only one study in which the advice given to teachers in their training, the general and specific do's and don'ts for instructional activities, was incorporated into an observational instrument, and an attempt was made to determine whether following this advice led to the expected student change (Kochendorfer, 1966). Even this single study of BSCS instruction was not followed up, and no attempt was made to use the results of this study to modify teacher training procedures and to determine whether such modifications led to enhanced gain in the students whom these teachers taught.

Of course, there are occasional studies on the relationship between general instructional activities and student growth within a curriculum program. Some of the research in Harvard Project Physics (Walberg, 1969) is an excellent example of research of this type. Unfortunately, such studies tend to be scattered, and the results of these studies are seldom used to revise the original curriculum package.

Specific observational systems have been developed and used as part of the developmental work in recent curriculum programs (such as IPI and PLAN) and in the products of regional educational laboratories (such as the Southwest Regional Laboratory and the Southwestern Cooperative Educational Laboratory), and one hopes that such work will be continued. The reports which I have read to date have been limited to descriptive studies and do not contain data on the relationship between the use of these specific activities and measures of student growth, or on whether teachers who used the program according to the intentions of the designers obtained higher student growth than the other teachers. There are a number of yet unpublished reports in this area which should be available shortly.

Thus, despite the mission-orientation of curriculum development programs, I have found few examples of systematic, rational refinement of these programs which was based on how these programs were used in the classrooms. Neither the curriculum development groups nor the funding agencies appear to recognize the need for instructional research at each phase of development and dissemination.

There is lack of instructional research within other developmental programs. A similar pattern exists, I believe, in the Office of Education's Bureau of Educational Personnel Development (BEPD) programs for elementary teacher education. These programs appear to be eminently suitable vehicles for instructional research on the effects of a teacher's instructional activities upon student growth. Yet the programs did not contain funding for instructional research; the research in these programs apparently is limited to determining how effectively teachers can be trained in certain skills. There is no funding within the BEPD programs for determining whether training teachers in the skills enhances student growth.

Such a pattern exists, I believe, in the regional educational laboratories. The emphasis there is upon development and dissemination, and these laboratories have succeeded in developing and disseminating curriculum packages and instructional procedures to a large number of teachers. Instructional research into how the materials are used and into the relationship between instructional activities and student growth has not received sufficient emphasis in the regional educational laboratories. Some research does indeed take place in these laboratories. For example, observational systems to monitor a teacher's use of materials have been developed in a number of laboratories. But these observational systems are seldom used in a systematic manner to develop new knowledge, determine which instructional behaviors should receive the greatest emphasis in the training program, or determine which behaviors are most important for enhancing student growth. Funded systematic research into instruction has not been built into the mission of these laboratories.

If the way in which an educational product is used is as important as the development of the product, then instructional research should take place as part of the development and dissemination work which occurs in the regional educational laboratories, teacher education programs, and curriculum development projects. I would hope that existing organizational and funding policies might be modified to permit this additional necessary research.

Summary

The current lack of knowledge on how teachers make a difference is not the result of poor research, or the lack of significant research findings. Rather, there has been a lack of research on how teachers make a difference. However, this lack of research can be remedied. In the next section, I shall present a description of some potentially useful research tools which have been developed, as well as tools which need development. The final section will contain a discussion on how these tools might be used in classroom research.

Tools for Future Educational Research

During the last 10 years, two educational inventions have been made which hold great promise for improving our ability to educate children in school settings: curriculum materials packages and systematic observational systems. Both developments are inventions in the sense that they have potential for helping teachers and researchers perform their tasks more efficiently and effectively. They can help teachers enhance the growth of their students and help researchers understand and contribute to the improvement of the instructional process.

These inventions are currently in a crude state of development, and we shall probably have to go through trial and error to learn how to use them. They are not panaceas, but they do offer promise for improving our ability to research, develop, and disseminate knowledge which will be useful to real children and real teachers in real schools.

The focus in this section will be on describing the current status of three tools or inventions—curriculum materials packages, systematic observation, and measures of student growth—and in suggesting how one might use observational systems and measures of student growth to conduct research on the relationship between instructional activities and student growth. The next section will focus upon conducting the research and analyzing the data.

Curriculum Materials Packages

During the 1960's we witnessed the development of numerous curriculum materials packages, such as the aforementioned BSCS program, the UISCM materials, the SMSG program, and IPI.

Although curriculum materials packages are prevalent today, this invention was probably developed independently numerous times in the past. The curriculum materials package produced about 1910 by Montessori (Evans, 1971) is a superb example. The Montessori Method included specific, self-correctional materials (the brown stair, the pink tower, the golden beads), specific instructions for teacher interactions with the child (vocabulary development with a three-period sequence: naming, recognition, and pronunciation), and general instructions for teachers (collaborative work with the child, avoidance of "don'ts," emotional support). The instructional materials, sequencing, provision for corrective feedback, and specific and general instructions to teachers contained in the Montessori materials are quite different from the usual practice of providing a teacher with only a set of books, a syllabus, and vague objectives. The major advantage of the Montessori package, or any curriculum materials package, is that it may enable a teacher to accomplish ends which could not be accomplished without these materials.

Curriculum materials packages have been developed by national curriculum groups, regional educational laboratories, universities, and private corporations from 1955 through the present. They represent a potential experimental treatment, and are analogous to the explaining experiment which Gage discussed (chapter 3) or the teacher training and feedback experiments which Flanders describes (chapter 6). Such packages share with the explaining materials and the teacher training materials the experimental possibility that when these materials are used, students will obtain desired ends more effectively than they could in classrooms which do not have these materials. The packages usually contain (a) the materials for use in a series of instructional units, (b) unit tests and remedial loops, (c) specific instructions to teachers on instructional activities, and (d) general instructions to teachers. The intention of the developers is to provide teachers and students with a sequence of instructional activities which are "packaged," that is, which have not been brought together in this way before. The packages are expected to contain the wisdom of the subject area experts, the practical knowledge of experienced teachers, and the instructional knowledge which has been developed in laboratory studies and small-scale classroom studies. Such packages of materials, instructions for the use of materials, suggestions for classroom activities, unit tests, and remedial loops are expected to have two additional advantages. First, the final package of materials and suggestions for instructional activities has been revised two or three times on the basis of reports of the teachers and measures of student growth. Second, there are teacher training sessions designed to help teachers learn how to use the materials, and such sessions are frequently supplemented by classroom visits by curriculum advisers. Thus, curriculum materials packages hold tremendous promise because they offer a systematic, rational procedure for improving teaching.

Curriculum packages and new curriculum programs which are developed and tested under a central agency are experimental programs. Teachers who are using these materials, particularly those who are following the intentions of the program designers, represent experimenters, who are replicating experiments across thousands of classrooms. The advantage of studying these teachers in situ is that we are studying them while they are using the materials and approaches which embody a specific program. Studies in such a context may be analogous to studying the behavior of physicians as they use new medical inventions and techniques.

Of course, the current state of this invention is crude. Many of the results of laboratory studies and small-scale classroom studies have not been incorporated into these curriculum materials packages. For example, the research and suggestions on sequencing which are discussed by Stolurow (chapter 7) have not yet been incorporated into curriculum materials packages. However, modifications can be made in the future to include more laboratory-developed variables in curriculum packages and to assess the effects of such modifications.

Perhaps the greatest deficiency of the current curriculum development model has been that the development of curriculum packages and the associated training of teachers has not insured that the teachers would use the package according to the intentions of the designers. As one researcher has noted:

> . . . those interested in curriculum development have not finished their job when they have packaged a cognitively valid and consistent set of materials. They must establish, in addition, how these materials are operationally introduced in the classroom environment. Otherwise, they will be left with

> unqualified assumptions as to how their package is unwrapped in the classroom (Gallagher, 1970, p. 102).

The importance of observing how a curriculum package is used in actual practice leads us to a discussion of the second invention, systematic observation.

Systematic Observation

A second invention that emerged during the 1960's (although, again, there are instances of this development in the past) is the development of techniques for systematic observation of instructional activities. These observational techniques represent an invention because they enable us to record quantitative and qualitative aspects of instructional activities in a way we could not do before.

The number and variety of these instruments are impressive. At least 200 instruments have been developed to enable an observer to record the frequency of specific events, and at least that many rating scales have been developed to estimate the frequency of occurrence of less specific events such as a teacher's clarity, enthusiasm, admission of error, and friendliness. Additional instruments, such as "sign systems," enable an observer to record whether or not an event occurred within a specified period such as 5 minutes, and sign systems can be used to observe both specific and general events. Complex observational instruments have been developed which enable one to record numerous facets of a single instructional event; for example, a teacher's question can be coded as to its cognitive level, the person addressed, the topic, the instructional intent, the tone, the number of students who listened, the appropriateness, and the subsequent student or class reaction. Recently developed rating scales not only include a large number of specific items, but also record whether or not a specific activity or behavior is congruent with the intentions of the curriculum developers.

In addition to observer instruments, new student questionnaires have been developed which allow students to state whether they agree or disagree that certain instructional activities are taking place, that the teacher tends to perform specific activities or encourage specific student activities, that the students enjoy their classwork, that students are dissatisfied, that the work is too difficult, that the class is well organized, or that the students are enthusiastic about what they are studying. In place of the older teacher questionnaires which focused upon the teacher's attitudes toward children, newer questionnaires focus upon specific classroom activities, the way the teacher responds to student comments, the classroom activities the teacher approves and disapproves, and the teacher's enthusiasm for various parts of an instructional program.

Such a burst of new observational instruments has undoubtedly led to misuse, misunderstanding, a lack of conceptual clarity in developing items, and more dogma than knowledge. However, despite these problems, observational instruments offer a systematic, rational procedure for improving teaching.

If we had a third invention we could take a giant step in the study of instruction. Unfortunately, this area seems so value laden that even a descriptive word for this invention raises hackles. But a third, necessary invention for the development of the rational, scientific study of education is a series of measures of student growth.

Measures of Student Growth

The term "measure of student growth" sounds awkward. It does not refer to height. I use it because it seems less offensive than synonyms such as "test," "achievement measure," "outcome," or "product." After all, "test" is a four-letter word.

There is a variety of outcomes of schooling about which the public and educational researchers would like to obtain information. Such outcomes might be classified into three types: academic achievement, attitudes, and personal development. "Achievement" refers to knowledge of facts, and also to skills of cognitive processing such as the ability to interpret, summarize, and compare information. "Attitudes" refers to a variety of measures which may or may not be interrelated: attitudes toward self, school, or subject areas; out of school activities, such as browsing in a library or going on nature walks; and dispositions to use cognitive skills in future activities. "Personal development" refers to a variety of outcomes such as self-confidence, ability to persist in difficult tasks, disposition to inquire into new problems, assumption of personal responsibility, ability to make reasoned choices, curiosity, and development of independence.

Unfortunately, it has been much easier to list and advocate objectives such as these than it has been to develop acceptable means of measuring their attainment. As long as we do not have acceptable means of testing outcomes, we shall be bound to the testimony of biased advocates of various programs. Testimony and social criticism are unquestionably important inputs for guiding educational research, development, and dissemination; however, testimony will not substitute for research and results in real classrooms.

Academic achievement is by far the outcome measure most acceptable to the majority of parents, students, teachers, and educators. Academic achievement was the sole outcome measure in the survey reported in *Equality of Educational Opportunity* (Coleman et al, 1966); it was the most notable outcome measure in the study of the effectiveness of Head Start programs (Cicirelli et al, 1969); and for several years major cities have been publishing annual mean achievement scores for their students in reading and mathematics. Current tests of reading comprehension are particularly well written and are measures of student ability to translate words in context, draw inferences from written material, and choose the most appropriate summaries of reading selections. New, updated editions of the Iowa Tests of Basic Skills, for example, reflect the current emphasis on social utility and relevance for a diverse population:

> Passages have been selected from a variety of current materials, including newspapers and magazines. . . . There are articles on conservation, pollution, and occupations. The majority of items require the pupil to draw inferences, make generalizations, and think critically (Hieronymus, 1971).

Reading comprehension remains an important educational end; the fact that many teachers prefer to spend class time on other activities, and the fact that some educators and social critics have not studied these tests do not detract from the importance which the American public attributes to reading comprehension.

But academic achievement is not the sole purpose of schooling today. Other educational ends are listed above. Although "we have reasonably good instruments for measuring academic skills and knowledge, we have essentially no capacity at all when it comes to measuring the aspects of personal development" (Mood, 1970, p. 8). For example, although student positive attitudes toward self, school, teacher, and subject area are desired educational outcomes, the research in this area is puzzling, at best. For example, of the seven or eight studies which I have seen in which student attitudes toward the teacher or the subject matter were measured twice—at the beginning and at the end of a school year or an instructional unit—student attitudes were always lower at the end of the instructional period (see Flanders, 1968; Gage et al, 1960). Thus, studies which attempt to relate instructional activities and student attitudes are, in effect, studies of instructional activities which are related to minimal loss in student attitudes. For example, Flanders and associates (1968) found that, in their sample, those teachers who praised students the most obtained less attitude loss than the other teachers in the sample.

But "minimal attitude loss" is hardly an acceptable outcome measure. The reason for this loss of positive attitude is difficult to determine. Perhaps the problem is the inadequacy of the measuring instruments.

Clearly, we need better inventions to measure the various educational ends which concern our society. This is a task of the highest priority, and I would recommend that all parity groups (students, teachers, parents, educators) be involved in developing these tests and assessing the suitability of various items. Currently, items tend to be retained or eliminated from tests on the basis of statistical considerations; other parity groups might wish to inspect tests in order to eliminate those vague and vexing items which contain more than one good choice, and to eliminate those items which appear to be unfair to various cultural subgroups.

As we approach a world in which more energy and time will be devoted to personal and interpersonal growth and less to productivity (Etzioni, 1971), measures which can assess each individual's personal development become more and more important. We are increasingly concerned about each student's social competence, self-confidence, ethics, and personal goals (Mood, 1970), bigotry or lack of bigotry (Sizer, 1971), dignity, autonomy, freedom, and range of human potential which has been developed. We are a long way from developing measures in these areas which can be confidently accepted; developing such tests is a much more difficult task than presenting testimony, nebulous hypotheses, or social criticism by the volume. Indeed, we are not using or refining those tests which we have. A seemingly excellent test on student sensitivity to social issues (Taba and McGuire, 1942) has lain dormant these 30 years.

The lack of suitable methods for measuring the effects of education disturbs me greatly because I am interested in grounding our intuitions in reason, logic, and empirical analysis. But I may represent a dwindling minority in education. Just as there are those who claim to know good art when they see it, we have an increasing number of educators who claim to know good education when they see events which match their poetic intuitions. Poetic intuitions, which once were considered a source of ideas for research, have increasingly become a sufficient end.

Summary

In summary, the development of acceptable methods for measuring educational change is the most critical task to be completed at this time. The currently crude inventions of curriculum materials packages and observational systems could be refined fairly easily if we had precise criteria against which to measure the refinements. Currently, the tests of academic achievement are excellent measures for assessing group progress, but they are limited as diagnostic tools for individual students. Currently, our testimony on student attitudes and personal development is far ahead of our technology.

Suggestions for Payoff Research in Real Classrooms

This final section contains suggestions for using the tools—curriculum materials packages, systematic observation, and measures of student growth—to conduct future research. The purpose of this research is to identify the ways in which teachers make a difference and to use the results of the research to modify instructional programs. Many of these suggestions follow from earlier papers by Cronback (1963) and Stake (1967).

The basic design for this research is presented in the list of "Questions for Instructional Research." Even researchers may not readily see how this design is applicable to a variety of situations; therefore, within this section, the same design is presented in four contexts: curriculum-specific studies with and without comparison groups, naturalistic studies, and teacher training programs.

Four points developed in the previous sections merit repetition. First, it was noted that the study of a situation in which a teacher is using a curriculum materials package is potentially more fruitful than one in which a teacher is using his "natural" skills. Situations involving curriculum materials packages are potential experimental treatments with which the teacher may be able to accomplish much more than he could without such aids. Second, it was noted that a curriculum materials package (or a teaching method, a structural rearrangement, or any innovation) is not a single variable. Rather, there is a good deal of variation in teacher behavior, student behavior, and instructional activities as a curriculum is being used. Third, the techniques of systematic observation offer the promise of quantifying variations in instructional behavior within the context of a specific curriculum package. Fourth, there are various outcome measures—measures of cognitive processing, attitudes, and personal development—which can be used as measures of student growth.

If the above four points are accepted, then we can conduct research to identify how teachers make a difference and to identify those features of programs which appear to be crucial in enhancing student growth. The basic research model consists of four sequential phases: (a) training teachers to use a certain package of materials, (b) studying the relationship between instructional activities and student growth *within* those groups of teachers who are supposed to be using the experimental treatment, (c) changing training procedures and/or materials on the basis of these studies, and (d) conducting new studies to determine the effects of the modifications and to determine the new relationships between instructional activities and student growth. None of these steps is unique or inventive, but the inclusion of all these steps as part of instructional research has seldom been done on a systematic, continuing basis.

Although the focus in this section is upon research on the implementation of curriculum materials packages, the ideas may be useful for any program which includes development and dissemination. Examples of such programs are open education, inquiry method, and preservice teacher training programs.

The major question in a study on how teachers make a difference is, what is the relationship between instructional activities and a variety of outcome measures. In order to answer this question, three types of data appear to be useful:

1. Measures of the occurrence of instructional activities considered important to the implementation of the specific curriculum.
2. Measures of the occurrence of instructional activities which have generally been related to outcome measures.
3. Measures of student growth on outcomes of interest.

Selecting Measures of Instructional Activities

It should be noted that there is no orthodoxy, no official list of measures of instructional activities which can be used to collect these data. There is a variety of instruments which can be used to obtain information on instructional activities, including student questionnaires, teacher questionnaires, observer questionnaires, and instruments to count the frequency of certain events or patterns of events. The investigator is free to choose or develop the instrument and items which are most appropriate to the research question. The fact that an instrument is useful for teacher training does not necessarily mean that it will be useful for a specific study.

In research of this type, measures of two overlapping sets of instructional activities may be useful. The first set is the behaviors which appear to be important for the implementation of the specific curriculum. The second set is the more general instructional activities which have been useful in past studies, and are of concern to educational researchers, participating teachers, and/or the audience that will read or support the study (Stake, 1970).

Clear, precise identification of the pertinent instructional behaviors in most curriculum materials packages is difficult because developers frequently are able to give only gross specifications. For example, in programs having a focus upon inquiry, the preferred instructional activities frequently include less teacher lecturing; more teacher questions which require higher thought processes, and more encouraging teacher responses; and more student talk, inquiry, and independent activity. To the extent that such variables are endorsed by a specific inquiry-type curriculum package, they would be included in the observational system.

In addition to rating and counting behavior in areas specific to the intentions of the curriculum developers, an investigator might also wish to collect information on instructional activities or teaching behaviors which are considered to be of general and pervasive importance in instruction. Such behaviors might include the clarity of the teacher's presentation, the methods used to correct or respond to incorrect student answers, and the enthusiasm of the teacher and the students. Again, this information might be obtained through rating scales, observational category systems, and/or student questionnaires. In the present state of the art, instructional research is not advanced enough to suggest items or areas which should be explored in any area of instruction. Some variables which appear worthy of future consideration have been summarized elsewhere (Rosenshine and Furst, 1971), but the reader should be cautioned that those variables were determined primarily through correlational and not experimental research, and they represent only those variables which were included in studies relating instructional behaviors to measures of student achievement. Other variables of equal, greater (or possibly, lesser) importance have yet to be studied, and reviews focusing on instructional activities related to noncognitive outcomes or personal development have yet to be written. Because of these deficiencies, an investigator is invited to explore additional variables of potentially general instructional merit.

An inexpensive observational instrument. Currently, it is possible to conduct a great deal of instructional research with very little expense by using student questionnaires, such as those developed by Kochendorfer (1966) or Walberg and Anderson (1969). In addition to providing information on teacher-student interactions, these instruments can be used to obtain descriptions of a variety of instructional activities which are not usually studied. These additional variables can include the amount and variety of homework assignments, the teacher's use of homework, the pacing of the classwork, the enthusiasm which the teacher and students have for the activities, the motivational procedures which are used and how the students respond to them, the students' assessment of whether specific topics or skills were taught, and whether the students felt they were treated as persons of worth. In those studies in which both student questionnaires and specific counts of instructional activities have been used, the questionnaires consistently yielded higher correlations with student change (see Rosenshine, 1970, pp. 286-87).

Investigators who have used these questionnaires have frequently mailed them to teachers across a wide geographic area, and the investigators have reported that the rate of return of the questionnaires was quite high. In future studies, the investigator might wish to submit the first draft of his questionnaire to a committee of students, teachers, and parents, and to develop the final document with them. Such a questionnaire might be more acceptable to everyone and might be a better instrument.

This emphasis upon questionnaires is not to be taken as a deemphasis of the importance of observer instruments which count the frequency of events. The optimal strategy for instructional research would be to use both types of instruments. The questionnaires can be used to identify instructional activities and teacher behaviors which appear to be important, and the counting instruments can be used to determine the specific behaviors which comprise such promising variables as teacher enthusiasm, clarity of presentation, or academic orientation of the teacher. All of the suggestions which Flanders (1970) and Furst and I (1971) have developed for the modification and refinement of counting instruments are still advocated. However, my current suggestion involves two steps: first, questionnaires are used to identify fairly gross instructional activities of merit, and second, counting instruments are used to identify the more specific behaviors.

Selecting Outcome Measures

It is particularly difficult to obtain suitable measures of the variety of outcomes of concern to program developers and the public. The lack of suitable instruments for measuring student change in attitudes and in areas of personal development has been discussed above. The problem becomes more difficult when we recognize that constructs such as achievement or self-concept are not single variables, and that tests which purport to measure reading comprehension do not always correlate highly with each other.

For now, one suggestion is that the investigator attempt to collect data on a variety of measures of student growth. Collecting data on both cognitive and noncognitive measures appears to be particularly important these days. An example of the use of a variety of outcome instruments is a study in which three cognitive and three noncognitive measures of student growth were obtained (Walberg, 1969).

Data Analysis

There is no orthodoxy for statistical analysis, other than the need to use the classroom, or subgroups within the classroom, as the unit of analysis instead of using each student as the statistical unit. Also, in research of this type, a scattergram is an extremely important tool for seeing how the data lie before one does more sophisticated analyses.

Once data on instructional activities and measures of student growth have been obtained, there are various ways in which they can be and have been analyzed. Ten possible procedures are listed in the "Questions for Instructional Research Within a Curriculum Package" which follows.

Questions for Instructional Research Within a Curriculum Materials Package

1. To what extent were the instructional activities within the program those which were intended by the curriculum developers?
2. Did the classrooms (or other units) within the program differ in their use of instructional activities specific to the program?
3. Did the classrooms within the program differ in the use of general instructional activities considered important for student growth?
4. Were the classrooms within the program different on the outcome measures of interest?
5. What was the relationship between use of program-specific activities and student growth?
6. What was the relationship between general instructional activities and student growth?
7. Were there differences in student growth among classrooms of teachers who were high, average, or below average in their fidelity to the intentions of the curriculum developers?
8. Were there differences in student growth among classrooms of teachers who were high, average, or below average in their use of general instructional activities?
9. Were classrooms which were high, average, or below average in student growth different in their fidelity to the intentions of the curriculum developers?
10. Were classrooms which were high, average, or below average in student growth different in their use of general instructional activities?

The first question focuses upon the extent to which there indeed was a set of variables which could be labeled a specific curriculum materials program. The next three questions focus on the extent to which the classrooms (or appropriate groups of students) were different on three measures of interest: specific instructional activities, general instructional activities, and measures of student growth. Frequently investigators do not bother to check the level of significance of the differences among the classrooms; they skip to questions 5 through 10, attempting to determine the relationship between instructional activities and student growth. However, in many studies, the finding of no significant relationship between instructional activities and student growth might be explained by checking the amount of variation among classrooms on the independent and dependent variables. (For an explanation of this problem, see Flanders 1970, chapter 12.)

The last six questions (5 through 10) represent three types of statistical analyses, each of which has been used in previous studies, which are useful in identifying the extent of the relationship between instructional activities and student growth. Three types of statistical analyses are suggested because in previous studies the same data have yielded widely discrepant probability levels depending upon the type of analysis. With three types of analyses (or a carefully studied scatterplot), one is in a position to describe the relationship between the instructional activities and student growth more clearly than one could by merely saying that the result was or was not statistically significant. Indeed, since I regard such correlational studies as useful primarily for developing hypotheses for experimental studies, certain levels of significance are not sacred; the replication of results across studies becomes much more important.

In questions 5 through 10, two types of variables are also suggested: those specific to the curriculum program, and general instructional variables. Although these two types overlap a great deal, some of the examples below may illustrate why they are separated at this time.

Question 7 focuses upon differential student growth when classrooms (or appropriate units) are grouped according to behaviors considered important to the implementation of the curriculum program. Such a grouping would seem particularly useful when curriculum developers have a general idea of the instructional activities which are appropriate to the program but cannot specify each behavior. A rating scale containing items referring to the teachers' fidelity to the intentions of the program developers would appear to be particularly useful in stratifying teachers. Once the teachers (with their classrooms) are grouped according to their fidelity to the intentions of the designers, the groups can be compared on student growth measures. If the student gains (on a variety of outcome measures) are higher in the fidelity group, subsequent questionnaires, observational forms, and counting instruments can be used to specify those behaviors and patterns of behaviors which also differentiate the groups. These data on instructional activities (and additional data which become necessary as a result of this analysis) could then be used in subsequent studies with the same sample (using correlational procedures or analysis of variance procedures) to obtain refined information about the importance of various instructional activities in enhancing student growth and/or about the range of such activities that is most functional. Presumably, as a result of both analyses, initial ideas would be modified, some ideas would be dropped, and new ideas would be developed.

As these two types of analysis are reiterated, certain instructional activities, patterns of activities, and ranges of these variables will likely emerge as important correlates of student growth. Such ideas can be investigated by incorporating them in the training programs for teachers. The effects of the modified training procedures could be studied, and modifications in training could again be made. The cycle would be continued until the curriculum materials package and the associated training program were satisfactorily effective in enhancing student gain on measures of interest.

One can easily think of refinements of these research suggestions. For example, one could use the same model to study whether certain instructional activities were more or less appropriate depending on the specific curriculum materials package, the readiness level of the students, the milieu of the school, and the home background and socioeconomic level of the students. One could also study whether the range of appropriate and inappropriate instructional activities for enhancing student growth varies for subgroups of students within classrooms. In addition, these research suggestions could be expanded to include comparison groups of teachers who have not received the training or the curriculum materials packages.

Examples of Research Within Curriculum Implementation

Curriculum-specific studies. As far as I can determine, there are almost no studies in which investigators attempted to relate the frequency of curriculum-relevant behaviors to measures of student growth. There have been isolated reports in which the investigators observed the curriculum-relevant behaviors of teachers (e.g., Olivero, undated; Katz, 1968; Niedermeyer and Dalrymple, 1970), but in reports such as these, the investigators did not attempt to relate the instructional activities to measures of student growth.

Of course, there have been a number of studies in which observers used a general observational instrument to relate teacher behaviors to student growth within the context of a specific curriculum program (LaShier, 1967; Soar, 1971; Walberg, 1969), and, hopefully, additional examples of such an approach are represented in research currently underway, particularly in the evaluation of the Follow Through model programs. Unfortunately, most of the observational instruments were designed to apply to all types of instruction. Such general instruments are important, but the exclusive use of general instruments has led to the neglect of variables which are important for a specific program. For example, I have yet to find an official, specific observational instrument written by the development staff of any of the national science programs. The observational instrument which was developed as part of the Oral Language Program of the Southwestern Cooperative Educational Laboratory (Olivero, undated) seems unique. I hope there are 20 similar instruments of which I am unaware.

One example of the research which I am proposing is a study in which the investigator developed a student questionnaire to determine the extent to which the teachers were following the practices recommended in the BSCS program (Kochendorfer, 1966). For scoring purposes, each item on the questionnaire was coded (by the investigator and judges) as to whether or not it represented the practices advocated as part of the BSCS program. In this study, there was a significant correlation (r = .30) between teacher use of the practices recommended by BSCS and student residual growth on the Processes of Science Test (BSCS, 1965). Unfortunately, as far as can be determined, there has been only one subsequent study in which this instrument was used to study BSCS teaching, and there has been no study to determine which of the many behaviors on this questionnaire appear to be the most important ones for student success in the BSCS curriculum. Without such knowledge, those educators who are training teachers for the BSCS curriculum are unable to modify and test their assumptions about which instructional activities are important for the BSCS curriculum.

Currently, Martin Siegel of the University of Illinois is the principal investigator in a pilot study of the type I have been describing and advocating. We are studying instructional activities and student achievement in the Distar program, a commercial model of the Bereiter-Englemann program. Siegel did not develop his observational system by looking at the 100 category systems in the *Mirrors for Behavior* volumes. Rather, he developed a series of rating scales based upon the teacher behaviors which were emphasized as do's and don'ts in the Distar teacher training program and in the meetings of supervisors. Several variables—particularly the extent to which the teacher requires a correct response from all the students—are yielding correlations of above .75 with the adjusted achievement scores. Having obtained significant correlations (and inspected the scatterplots), Siegel is now developing category systems to count events within those variables which were most highly correlated with student growth. Once the category systems are developed, we shall hire raters to count the frequency of events on the audiotapes. In addition, we are now in a position to rate or count other variables of general interest which appear on the audiotapes; we are planning another study, one involving 30 teachers and including noncognitive measures, for next year. Even now, Siegel is applying the strongest findings to the training of teachers for the Distar program.

An important aspect of Siegel's research was obtaining a significant correlation between teacher behavior and student achievement in a program such as Distar in which the teacher behavior is highly prescribed. Such variation in teacher behavior and student achievement suggests that even a highly specified program such as Distar cannot be considered a single variable.

The addition of comparison groups. There is the possibility that a given curriculum materials package is not particularly distinctive, and that the results obtained through using this curriculum are not improvements over traditional instruction. This possibility should be tested, and such testing usually takes place by comparing the mean growth scores in one program with the mean growth scores in the other program(s).

However, in any situation in which a curriculum package is not a single variable (and the traditional program cannot be considered a single variable either), an analysis which is limited to mere comparison of growth measures between two programs does not make use of the available information on how the teachers in the special curriculum and in the traditional instruction make a difference.

The comparison of outcome measures alone overlooks the knowledge to be gained from also determining relationships between instructional behaviors and student growth both within the special curriculum and within the comparison program. Information on instructional activities within each program is not particularly difficult or expensive to obtain; student questionnaires which would take 30 minutes of the students' time could be used to obtain student perceptions on whether or not specific instructional events took place, on what responses the teacher made to various student actions, on interactions among classmates, and on other variables of interest. If a study comparing the outcomes of two programs is supplemented by data on instructional activities, then we can conduct three analyses: (1) comparison of the two programs on outcome measures, (2) determination of relationships between instructional activities and student growth within the special curriculum group, and (3) determination of the relationships between instructional activities and student growth within the traditional group.

The value of probing for instructional relationships within the experimental classrooms (or appropriate unit) and also within the control classrooms becomes particularly apparent in studies in which there were no significant differences between the two groups on measures of student growth. In such a situation, the term "no significant differences" usually means that there was wide variation within each group on the outcome measure. Given such wide variation, it seems particularly important to learn what instructional activities were related to this variation. Although there have been several studies comparing student growth in one situation with student growth in another (and although many of these investigators have also gathered data on instructional activities), none of the investigators, to my knowledge, has completed the study by relating instructional activities to student growth within each curriculum or situation.

Naturalistic studies. The emphasis in this chapter has been on the potentially greater payoff from instructional research conducted while all the teachers are using the same curriculum materials package compared with the payoff from observing teachers when they are not using such materials. The usefulness of studying the implementation of materials packages is that one is able to introduce new teaching behaviors and instructional activities which are unlikely to occur in the natural situation. But this emphasis should not be taken as an argument against continuing naturalistic studies; there is a great deal to be learned from continuing such research.

By naturalistic studies, I do not mean descriptive studies of instructional activities. Description alone has been well covered. I do refer to studies which attempt to relate the natural instructional activities of students and teachers to measures of student growth. In such naturalistic studies, two of the three tools described in the second section—observational systems and measures of student growth—are used as rigorously as possible; only the curriculum materials package is omitted. Detailed suggestions for conducting such research are presented elsewhere (Rosenshine and Furst, 1971).

One potentially fruitful area for naturalistic studies is contained in the proposal by Dyer (1970) in which he hopes to identify schools with children of similar initial ability and socioeconomic status, but in which the subsequent growth of children in an area such as reading is quite varied.

Once such schools have been identified, two types of research activities can be conducted. One study would focus upon studying those schools which were above average, average, or below average with respect to reading achievement, for example. All of the suggestions on the use of general observational instruments could be employed to determine how the instructional activities in the above-average schools (or classrooms, or appropriate units) differed from those in the average and below-average schools. However, whatever results are obtained in such studies, they represent correlates, not causal variables. The validity of these variables as causes of enhanced reading achievement could be tested in a subsequent study in which teachers in the below-average schools (or teachers in below-average classrooms) were trained to use the instructional activities which appear to exemplify the teachers in the above-average classrooms. To the extent that such training resulted in desired instructional activities and enhanced reading comprehension, we would be obligated to continue our search for variables which illustrate how teachers make a difference.

Instructional research in a teacher education program. Many of the above suggestions for instructional research can also be applied to a teacher training program, in which teachers are taught to perform a variety of instructional activities. After teacher training is completed, at least two important questions remain: (1) Are the teachers using the behaviors or activities in new situations? and (2) Does the use of the behaviors lead to improved student growth? One method for testing both questions would be to require preservice teachers to teach special units to groups of students. The special instructional units could be selected from various units that have been used in similar studies. Such assembled units contain both the materials and criterion tests whose internal consistency has been developed to a satisfactorily high level.

The teachers would teach these materials to students; during and after the lessons, observer forms, student questionnaires, and counting instruments could be used to obtain information on instructional behaviors. Measures of student change would be obtained on variables of interest.

In effect, data would be gathered in three areas: instructional activities which were taught in the teacher education program, general instructional activities which were not specifically taught, and student growth in areas of interest. These data are quite similar to the type described above in studies of curriculum packages, and the potential analyses are identical to those in the "Questions for Instructional Research." Only one study was found which used such a procedure, that conducted by Baker (1969).

In any analysis conducted within a group of trained teachers, there remain two difficulties: first, it is possible that the teachers were so well trained that there were no significant differences among the teachers on instructional activities of interest (questions 2 and 3); second, it is possible that even though student growth occurred, it was so evenly distributed that there were no significant differences among the classes (question 4).

Both hypotheses could be tested by obtaining data also from a control group of teachers who had no training, or who had no training in the specific program. If comparisons were made between the behaviors of teachers in the trained and the untrained groups, then the data could be used to answer each of the rival hypotheses. Such analyses are identical to those advocated above in the discussion of instructional activities of teachers in the special curriculum group and in the comparison group.

When researchers conduct a study on the effects of teacher training and collect observational data and growth data on both the experimental and control groups, they then are in a position to conduct at least four analyses:

1. Comparison of the mean growth scores in the two groups.
2. Comparison of the differences in instructional behaviors for the two groups of teachers.
3. Determination of the relationship between instructional activities and student growth in the experimental group.
4. Determination of the relationship between instructional activities and student growth in the control group.

Most of the studies in the past were restricted to the first analysis, the comparison of measures of gain. Recently, several investigators (Carline, 1969; Davidson, 1967, 1968; Herman et al, 1969; Millett, 1969; Rogers and Davis, 1970) have collected data on both the outcome measures and the behavior of the teachers. However, none of these investigators has used the data which he already had to compute the relationship between instructional activities and student achievement within each of the two groups. The larger reports on these studies do not contain sufficient data to allow an independent investigator to determine the within group relationships. Such waste is unfortunate, because almost all the work has already been done, yet the important data on the relationship between instructional behaviors and student growth are lost to the educational public. Hopefully, future researchers and supervisors of research will be alert to the additional analyses which can be conducted in such situations.

Summary

We are currently in a position to learn a great deal about how teachers and teaching make a difference. Educational products such as curriculum materials packages have the potential to enable a teacher to accomplish much more than he could without the materials and suggestions for their use; observational systems enable us to obtain information on how the curriculum was used and to identify those features which appear most promising; hopefully, there will be an increased use of tests and test situations which will enable us to measure a variety of educational goals. Hopefully, we shall begin to use these instruments for the systematic, rational study of teaching, to learn not only how teachers make a difference, but how to help teachers make a difference.

Acknowledgments

Norma Furst, Lilian Katz, Karin Myers, and Barbara Rosenshine provided helpful comments on an earlier draft of this paper.

References

Baker, E. L. "Relationship Between Learner Achievement and Instructional Principles Stressed During Teacher Preparation." *Journal of Educational Research*, 63, 3, 1969, pp. 99-102.

Biological Sciences Curriculum Study. *Processes of Science Test Manual*. New York: The Psychological Corporation, 1965.

Bond, G. L. and Dykstra, R. "The Cooperative Research Program in First-Grade Reading Instruction." *Reading Research Quarterly*, 2, 1967, pp. 1-142.

Carline, J. L. "An Investigation of the Relationship Between Various Verbal Strategies of Teaching Behavior and Achievement of Elementary School Children." Unpublished doctoral dissertation, Syracuse University, 1969. Also available as a paper presented to the American Educational Research Association, March 1970. Morgantown: West Virginia University, Division of Education.

Cicirelli, V. *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development*. Washington, D.C.: Office of Research, Planning, and Program Evaluation, Office of Economic Opportunity, 1969.

Coleman, J. S., Campbell, E. Q., et al. *Equality of Educational Opportunity*. U.S. Office of Education. Washington, D.C.: U.S. Government Printing Office, 1966 (OE-38001).

Cronbach, L. "Evaluation for Course Improvement." *Teachers College Record*, 64, 1963, pp. 672-83.

Davidson, R. L. "The Effects of an Interaction Analysis System on the Development of Critical Reading in Elementary School Children." Unpublished doctoral dissertation, University of Denver, 1967.
See also *Classroom Interaction Newsletter*, May 1968, pp. 12-13.

Dyer, H. "Toward Objective Criteria for Professional Accountability in the Schools of New York City." *Phi Delta Kappan*, 1970, pp. 206-11.

Etzioni, A. Essay review on *Crisis in the Classroom* by Charles A. Silberman, *Harvard Educational Review*, 1971, 41, 1, pp. 87-98.

Evans, E. D. *Contemporary Influences in Early Childhood Education*. New York: Holt, Rinehart and Winston, Inc., 1971.

Flanders, N. A. *Analyzing Teaching Behavior*. Reading, Mass.: Addison-Wesley, 1970.

______, Morrison, B. M., and Brode, E. L. "Changes in Pupil Attitudes During the School Year." *Journal of Educational Psychology*, 50, 1968, pp. 334-38.

Gage, N. L., Runkel, P. J., and Chatterjee, B. B. *Equilibrium Theory and Behavior Change: An Experiment in Feedback from Pupils to Teachers*. Urbana: Bureau of Educational Research, University of Illinois, 1960.

Gallagher, J. J. *Teacher Variation in Concept Presentation in BSCS Curriculum Program*. Urbana: Institute for Research on Exceptional Children, University of Illinois, 1966 (ERIC: ED 023 206).

______. "Three Studies of the Classroom." In *Classroom Observation*. Chicago: Rand McNally and Company, 1970. Also in Gallagher, J. J. *Analyses of Teacher Classroom Strategies Associated with Student Cognitive and Affective Performance*. Cooperative Research Project No. 3325. Urbana: University of Illinois, 1968, p. 43 (ERIC: ED 021 808).

Herman, W. L., Potterfield, J. E., Dayton, C. M., and Amershek, K. G. "The Relationship of Teacher-Centered Activities and Pupil-Centered Activities to Pupil Achievement and Interest in 18 Fifth-Grade Social Studies Classes." *American Educational Research Journal*, 1969, 6, pp. 227-40.

Hieronymus, A. N. Interview in *Bureau of Educational Research Pipeline*. New York City Public Schools, 1971.

Katz, L. G. "Children and Teachers in Two Types of Head Start Classes." *Young Children*, 24, 1968, pp. 342-49.

Kochendorfer, L. H. "A Comparative Study of the Classroom Practices and Teaching Rationale of High School Biology Teachers Using Different Curriculum Materials." Unpublished Ph.D. dissertation, The University of Texas at Austin, 1966.

LaShier, W. S., Jr., and Westmeyer, P. "The Use of Interaction Analyses in BSCS Laboratory Block Classrooms." *Journal of Teacher Education*, 18, 1967, pp. 439-46.

Lindvall, C. M. and Cox, R. C. *The IPI Evaluation Program: AERA Evaluation Monograph No. 5*. Chicago: Rand McNally, 1970.

Millett, G. "Comparison of Four Teacher Training Procedures in Achieving Teacher and Pupil Translation Behaviors in Secondary School Social Studies." Paper presented to the American Education Association, 1969. Austin: University of Texas (ERIC: ED 027 256).

Mood, A. M. "Do Teachers Make a Difference?" In *Do Teachers Make a Difference?* U.S. Office of Education. Washington, D.C.; U.S. Government Printing Office, 1970 (OE-58042).

Niedermeyer, F. C. and Dalrymple, M. C. "The Development and Evaluation of an Exportable Teacher Training Package for a

Kindergarten Reading Curriculum." Paper presented to the American Educational Research Association, February 1970. Los Angeles: Southwest Regional Laboratory for Educational Research and Development.

Olivero, J. L. *Developing the Oral Language Program*. Albuquerque, N. Mex.: Southwestern Cooperative Educational Laboratory, undated.

Rosenshine, B. "Evaluation of Classroom Instruction." *Review of Educational Research*, 40, 1970, pp. 279-301.

______, and Furst, N. "Research on Teacher Performance Criteria." In B. O. Smith (ed.) *Research on Teacher Education: A Symposium*. Englewood Cliffs, N.J.: Prentice-Hall, 1971, pp. 37-72.

Rogers, V. M., and Davis, O. L. "Varying the Cognitive Level of Classroom Questions: An Analysis of Student Teachers' Questions and Pupil Achievement in Elementary Social Studies." Paper presented to the American Educational Research Association, March 1970. Lexington: University of Kentucky.

Sizer, T. R. "Testing: Americans' Comfortable Panacea." In *Proceedings of the 1970 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service, 1970.

Soar, R. S., Soar, R. M., and Ragosta, M. "The Validation of an Observation System for Classroom Management." Paper presented to the American Educational Research Association, February 1971. Gainesville: University of Florida.

Stake, R. E. "The Countenance of Educational Evaluation." *Teachers College Record,* 68, 7, 1967, pp. 523-40.

_________. "The Decision: Does Classroom Observation Belong in an Evaluation Plan?" In *Classroom Observation*. Chicago: Rand McNally and Company, 1970, pp. 1-6.

Taba, H. and McGuire, C. "Evaluation of Social Sensitivity." In Smith and Tyler (eds.) *Adventure in American Education, Volume Three*. New York: McGraw-Hill Book Company, Inc., 1942, pp. 157-244.

Walberg, H. J. "Predicting Class Learning: An Approach to the Class as a Social System." *American Educational Research Journal*, 6, 1969, pp. 529-42.

_________, and Anderson, G. J. *Learning Environment Inventory.* Cambridge, Mass.: Harvard Project Physics, 1967.

Chapter 6

A NATIONAL COORDINATED PROGRAM OF RESEARCH ON TEACHING EFFECTIVENESS

Ned A. Flanders

Introduction

As long as children are required by law to come into contact with teachers, the analysis of these contacts with an eye to improving them is an inescapable, moral obligation of professional educators. Improving the contacts a child has with a teacher is a very complex problem, on a national scale it will be costly and will take a long time. A coordinated, cooperative, nationwide effort will require leadership and support from the Office of Education. It would appear that attempting to understand how a teacher influences educational outcomes by studying his interaction with students is essentially a Federal activity, smaller units of government just don't have the resources. The size of the task becomes easier to comprehend by imagining the steps we would carry out.

Step 1: We must try out systems for defining and describing the interactive events of teacher and learner contacts. In order to understand these interactive events, we may also wish to study those events which occurred as preparation and those that comprise evaluation.

Step 2: We must clarify our concepts for thinking about educational outcomes and sharpen our procedures for evaluating these outcomes. In particular, I suggest learning how to assess (a) pupil perceptions of teaching behavior; (b) pupil attitudes toward the school, toward the teacher, and toward one's self as a learner; (c) pupil behavior while at work, especially in terms of self-direction and independence; and (d) progress in subject matter achievement, that is, the substantive knowledge, skills, and attitudes which are the objectives of teaching.

Step 3: Complete preliminary field tests in order to identify promising associations between interactive events and educational outcomes. There are many possible associations. One example is the association between learning facts and variables based on assessing drill activities. Another example is the association between developing positive pupil attitudes toward the class and having the teacher respond to ideas expressed by pupils.

Step 4: Complete preliminary field tests to find out if experienced teachers and college students can learn to create particular interactive patterns by controlling their own teaching behavior. For example, can teachers be taught to respond constructively to ideas which pupils express?

Step 5: Repeat steps 3 and 4 with the format of a pure experiment so that variation of the principal variables can be guaranteed, conditions are more carefully controlled, and the integrity of the intended behavior is more adequately developed.

Step 6: The way "to put it all together" is to conduct pure experiments to find out if changes in teaching behavior actually cause changes in the educational outcomes. For example, can teachers learn to act differently so that the intended change improves the positive attitudes which pupils have toward their schoolwork? There are many different associations which need to be tested.

It is most encouraging to note that considerable progress has been made, during the past decade, at steps 1, 3, and 4. We do have systems for describing interactive events, we have used these systems in field studies on teacher effectiveness, and we have completed field studies in which teaching behavior was modified. At no time in our history have we been so well prepared to continue working on these steps. The challenge, then, is to build on present progress and increase, rather than decrease, our research momentum. During the decade of the seventies, we will have "to play catch-up" at step 2 because we are behind in techniques for assessing outcomes. Nevertheless, I believe we are prepared to begin a national effort at step 5 and step 2 simultaneously.

Overview

This paper is a proposal for a *National Coordinated Program of Research on Teaching Effectiveness* to collect raw data about the interactive events between teachers and learners. The data will include (a) magnetic video/sound recordings of the interactive events; (b) measures of pupil behavior, abilities, perceptions, attitudes, and subject matter learning; (c) measures of selected perceptions of the teachers; and (d) data describing the conditions that existed when the learning occurred. It is further proposed that these data be cataloged and stored in at least one *National Cooperative Archives for Recorded Teaching and Learning Behavior*. It is anticipated that the national archives will be a center to which scholars could come in order to try out various schemes for abstracting and quantifying the recorded interactive events. They then could find out which variables are, in fact, associated with desirable educational outcomes.

The data accepted by the national archives would have to be collected under conditions which satisfy certain minimal levels of experimental control. Among these conditions I would propose the following: (a) a prelearning measure of general academic ability; (b) a prelearning measure of performance on the learning objectives; (c) a standardized unit of instruction with specified instructional materials; (d) a controlled schedule of time devoted to learning; (e) planned, practiced, and verified variation in the behavior of the teachers; (f) appropriate postlearning measures of attitudes and achievement; and (g) the stratified, random assignment of students to experimental treatments so as to make between-treatment differences on student prelearning (or attitude) a minimal, chance occurrence.

The most important product of this program would be the video/magnetic recordings themselves, the raw behavioral data that could be analyzed and reanalyzed by different, competing systems. Soliciting and producing additional recordings would be a continuing affair as different settings, different teaching objectives, and different students are sampled, each with sufficient replications so that cross validation of results would be possible. Thus, different settings might highlight such comparisons as rural versus urban, inner city versus outer city, open classrooms versus more structured classrooms, and so on. Among the comparisons of teaching objectives we would expect to find learning particular content versus learning whatever the student wanted to learn, seeking independence and self-direction as a measure of learning rather than content, learning skills and principles rather than facts, and so on. The variations among students could include homogenous and heterogenous racial mixtures, deprived versus advantaged, each comparison to be made at all age ranges from kindergarten to undergraduate college, and so on.

Besides the production of raw data, other useful products can be expected. For example, we can expect (a) that the video/sound recordings would lend themselves to the production of instructional materials in the field of teacher education; (b) that standardized, evaluative, units of instruction would become available which could be used to compare alternative methods of teacher education—that is, the units could be used to evaluate the performance of beginning teachers; (c) that certain basic patterns of teacher questioning, teacher responding, and student initiating are likely to emerge as critical skills of instruction; (d) some notions about how teachers might modify their teaching behavior would be identified; and (e) that certain kinds of simulated, interactive experiences would be identified that show promise of helping college students develop critical teaching skills. These "secondary products" would represent a substantial contribution in their own right.

Justification and Limitations

How *do* teachers make a difference? How *can* teachers make a difference? These questions make use of concepts that have unclear meanings. Given this ambiguity, we are not likely to agree about the answers. We have to decide just what we mean by the questions and then seek to understand the different answers. I am reminded of Kaplan's[1] comment—

> In the scientific community, communication is also a matter of goodwill, the wish to understand. Scientists engage not only in mutual criticism, but also in mutual support.

One place at which controversy will arise over this paper is what might be called the narrow focus on overt behavior during teaching and learning. I can understand this criticism, but I would point out that by starting with the interaction which occurs during the contacts which teachers have with pupils, we are seeking a point of departure. Let me illustrate by making an oversimplification of the problem. One can start with values or principles of pedagogy and then ask what teaching behaviors then follow. One can also start with behaviors and then ask what values and principles do these behaviors illustrate. In either case, whatever we find out about teaching effectiveness must be translated into behavior, sooner or later. To start with behavior at least gives it a central position in our thinking.

There will also be controversy about the concepts which we choose as tools for analysis. The choice *is* difficult. As Kaplan[2] points out—

> The appropriate conceptualization of a problem already prefigures its solution.
>
> —and then adds the wry observation,
>
> we are caught up in a paradox. . . . The proper concepts are needed to formulate a good theory, but we need a good theory to arrive at the proper concepts.

Under these circumstances progress is made by inching forward on all fronts so that better concepts and better theories gradually emerge.

Through conjecture and argument, we identify starting points. The skill here is what Thelen[3] calls "the instinct for the jugular" and what Mae Brodbeck[4] meant when she wrote, "Some features of the world *stand out,* almost begging for names." My instinct, when considering how teachers might make a difference, is to look at the point of contact, at the transactions between teacher and learner. I chose to study and conceptualize the interaction which occurs when learning and teaching are taking place. For me, these are the phenomena that stand out, begging for names. My assumption is that more effective teachers interact differently than less effective teachers.

Unfortunately, current teaching practices in the field may be a poor resource for analyzing the full range of quality in teacher and learner contacts. One way to create a wider range is to plan it, practice creating it, and with increasing skill actually produce it. The objective is that a wide range of interaction patterns are created under controlled conditions such that pupils will perceive these various patterns as being authentic, genuine, and possessing integrity.

Notice that this approach does not restrict our study to classrooms which may go out of style as other instructional alternatives become available. Interaction between learner and teacher can take place on a one-to-one basis, in small group settings, in the classroom, as well as in massed audience formations. We also need not be restricted to interaction which involves the teacher, even though we may start at that point. For example, a skillful teacher may create opportunities in which students interact with other students. How such opportunities are "set up," then, may also be an object of study.

Another point to keep in mind is that we prefer concepts which help us in terms of our purpose. At this conference, our purpose is to find out how teachers make a difference. Concepts which don't have direct connections with a teacher or his behavior are less likely to satisfy this purpose. Thus, even though we know that the best predictor of subject matter achievement is a student's earlier performance on the same test, it is difficult to see how this information will help a teacher improve except to suggest that he avoid contacting below average students.

By way of summary, the concepts which will be of central interest fall into three broad categories. First, we will need concepts which differentiate the qualities of the contacts between teachers and learners. Second, we need concepts which differentiate among educational outcomes. And third, we need concepts which will help us identify features of a learning situation which must be controlled in order to study associations between interaction variables and outcome variables. These concepts will become tools for fashioning knowledge about what a teacher does that makes a difference. In turn, this kind of knowledge has direct application to teacher education at both the preservice and inservice levels.

A National Plan to Coordinate Research on Teaching Effectiveness

There are two parts to this plan: first, the National Cooperative Archives for Recorded Teaching and Learning Behavior; and second, a National Coordinated Program of Research on Teaching Effectiveness. In this section some features of both parts of the plan are discussed. What is written here is by no means a complete plan, nor even a complete outline. Making such a plan is a much more ambitious project than writing this paper.

Consider for a moment the six steps outlined on the first pages of this paper. We have now made enough progress with uncoordinated projects to be able to visualize those six steps and to identify some research tools which will facilitate the research required at each step. For example, we have access to more than 150 different systems for coding interactive behavior[5] and some progress has been made in approaching the problems of observer reliability and training. We can now recommend certain procedures of test construction in order to evaluate teaching and learning behavior. We can recommend recording techniques and equipment and we can recommend certain minimum standards of research design. It is the achievement of this much progress that makes national coordination a practical alternative.

The National Cooperative Archives

The archives is a physical facility something like a research center at which recorded teaching and learning behavior is stored. Since the main collection will consist of magnetic recordings and associated data printouts, it is obvious that more than one center can be created at the relatively small cost of duplicating the basic data. Besides serving as a storage site, there are other interesting functions which might occur at an archives center.

Data Bank Function: The archives will have many features in common with a library including cataloging, storage, retrieval, and acquisitions. Cataloging will start by listing the obvious characteristics of a recording such as grade level, subject matter, racial mixture, socioeconomic level, and so on. However, cataloging will be a continuing process not only because of new acquisitions, but because study and analysis by archive staff members will reveal additional information which needs to be catalogued. For example, teaching patterns that are more effective in terms of some educational outcomes can be distinguished from those which are less effective, situations in which pupil attitudes are more positive can be separated from the less positive, and so on. Encoding and analysis will also identify the presence or absence of particular behavioral events, thus, the catalogue could indicate those recordings in which the use of teacher praise is above average. Nearly any feature of interaction can be catalogued so that one problem will be to develop a useful classification scheme.

Storage and retrieval will require playback carrels with voice and videotape recorders, modern editing and duplicating equipment, and the more common data analysis equipment.

The acquisitions department of the archives would be responsible for organizing research contracts and support funds which, in effect, would consist of Requests for Proposals (RFPs). An RFP would be issued specifying certain characteristics of teaching and learning and identifying certain research design considerations so that acquisitions match the priorities of the research program. For example, an RFP might specify a request for recorded behavior at the fourth grade, inner city, with tutorial interaction between an individual pupil and a teaching aide in the field of remedial arithmetic. In addition, the recordings must sample pupils, aides, and interaction from two different patterns of tutoring so that inferences can be reached concerning the relative effectiveness of the two patterns. All of the interaction may be on video plus sound or some of the recordings may be on sound only. It is easy to imagine the nearly infinite number of combinations which would determine the specifications for the RFPs; probably some kind of a board of directors would be needed to set priorities and establish policies.

Encoding and Decoding Behavioral Events: One great advantage of an archive center is that permanent equipment can be installed to facilitate encoding and decoding. It has already been demonstrated[6] that a computer based system can create, in effect, a coding factory at which all collating and data analysis is automated from the instant that an observation button is depressed. The presence of an event or the assignment of a rating is electrically stored in the computer and then reproduced in a preprogramed printout display. It is also possible to visualize the encoding of behavior in live classrooms, using telephone transmission lines, connected directly to the computer used by an archives center. Preliminary work[7] for this has also been completed. Such a function may prove valuable when the performance of a particular teacher is to be compared with performance data already stored.

Permanent equipment for coding teaching and learning behavior can follow a flexible design so that it can be used with practically any observation system. A main purpose of the archive would be to encourage the analysis of behavior with different systems in such a way that comparisons between systems can be made. Electric pushbuttons and computerization provide flexible coding arrangements to fit different systems, to analyze errors in coding, and for the indepth study of coding reliability.

The Resident Staff and Visiting Fellows: The resident staff members of an archives center would carry out their own research, participate in the development of instructional products for teacher education, and act as consultants to visiting fellows. Predoctoral and postdoctoral fellowships woud permit scholars to carry out various projects concerned with teaching effectiveness.

There is no need here to give a long list of examples of the kinds of projects which could take place at an archive center. We should note, however, that these research and development activities are more practical and applied than they are basic research. We study associations between educational outcomes and different interaction patterns because we must make decisions about teacher education and select approaches to the improvement of instruction. It matters little whether the project makes use of a system for analyzing the cognitive aspects of discussing controversial social issues in secondary schools[8] or compares two groups of teachers who have been exposed to teacher training programs, the research is directed toward practical decisions about teaching effectiveness.

Teacher Education: At the present time we have various plans for different teacher education programs, but relatively few opportunities to develop instructional materials which have a proven capability for helping a trainee learn and perform a particular teaching skill.[9] Permanent or visiting staff at an archive for recorded teaching behavior could carry out the following steps if they wished. First, they could identify particular skills which are known to be associated with desired educational outcomes. Second, they could study repeated instances of how this skill is used. Third, they could design and develop some instructional materials for the purpose of teaching teachers the skill in question. Fourth, after they had tried out these instructional materials, they could compare the performance of their trainees with the recorded behavior of teachers whose data were in the archive and who were classified as very effective with regard to achieving the desired educational outcome.

Changing Teaching Functions: As CAI, IPI, and various forms of automation come of age, we are expecting changes in the role of the teacher and in the teaching station that he mans. Early research on interaction analysis would suggest that lecturing and asking simple questions, giving directions, and conducting recitation or drill occur with high frequency and are easy for teachers to perform. It is precisely these simple expository and single step questioning techniques which are most likely to be automated. Should these predictions come to pass, teachers will need assistance in developing skills like asking higher order questions, responding to ideas which pupils express, learning how to help pupils summarize their own agreements, and similar skills. Since these "responsive" skills, to give them a name, are currently used less frequently than the expository skills, when automation finally comes, it will force teachers to develop their responsive skills and to use these skills more frequently.

An archives center will be concerned with the changing role of the teacher in several ways. First, research on teaching effectiveness is necessary to identify critical responsive skills and then to find out whether these skills are associated with particular educational outcomes. For example, earlier research[10] would suggest that teacher responses to pupil ideas are significantly associated with positive pupil attitudes. Second, do these skills need to be modified as the proportion of automated instruction increases? So far, we have relatively little hard data on critical teaching skills which are necessary to support a curriculum involving CAI, CMI, IPI, and other innovations.

These and similar problems could be investigated at an archives center.

Summary: A National Cooperative Archive for Recorded Teaching and Learning Behavior would be one or more centers at which the recorded contacts between teachers and students could be stored and analyzed. We would expect it to produce knowledge about teaching effectiveness which applies to a wide variety of settings, to many different kinds of teachers, and to the various kinds of boys and girls to be found in our schools. The staff of such a center could investigate those functions that a teacher provides when most instruction is automated by several different methods. An archive center would be the focal point of any coordinated program of research on teaching effectiveness along the lines discussed in the next section.

Coordinated Research On Teaching Effectiveness

Field studies on teaching effectiveness, during the last decade, often reveal two disturbing flaws. First, the researcher did not develop procedures for assessing the particular learning objectives which the teacher intended to teach. Often the compromise consisted of using a nationally standardized test covering subject matter that the teacher did not intend to teach. Except for subject matter achievement, other outcome variables were usually neglected. Second, the range of teaching behavior has been too narrow. This, in turn, curtails the range of interaction patterns and limits the qualities which otherwise might be present in innovative teacher-pupil contacts. Much of the material in this section is concerned with improving these two aspects of research on teaching effectiveness.

A National Coordinated Program of Research on Teaching Effectiveness means that Federal funds would be allocated with specific restrictions. Restrictions are less onerous, however, when they are minimal requirements concerning generally accepted standards for replicable research. These minimal standards would concern the research design and the need to obtain intelligible video/sound recordings.

Recommended Research Designs: A first requirement is that pupils are assigned to learning sessions in such a way that the groups to be compared are approximately equal in terms of initial ability. A generalized test of academic ability and a specific pretest of the objectives of instruction are minimal requirements. With knowledge of initial ability, pupils can be assigned to learning groups by some form of stratified, random assignment. Given the administrative problems of our public schools, most experiments for this program may have to be organized during the summer or some other vacation period. Alternatively, there may be some settings in which pupil assignment would be possible during the regularly scheduled school year.

A second requirement would be that evaluative teaching units be developed for the experimental teacher's use. An evaluative teaching unit is discussed on subsequent pages, we note here that such units provide control of the instructional materials, the amount of time devoted to instruction, and the scheduling of that instruction.

A third requirement is that patterns of teaching behavior be planned, practiced, tried out, refined, and verified prior to the experimental teaching and learning sessions. Since we are interested in creating a wide range of interaction when teachers come into contact with pupils, this aspect of the research should be planned with great care. For example, it might take 6 to 9 months to locate teachers and then train them to perform certain teaching skills. With practice, experimental teachers can learn to provide fairly consistent patterns of interaction and hopefully the pupils would see these patterns as believeable and relevant. At the same time, between-teacher differences should be large so that the necessary wide range of behavior is present.

A fourth requirement is that intelligible sound/video recordings of the contacts between teachers and pupils be made. This requirement may result in somewhat smaller class size, accoustically treated rooms, and a period of adjustment for the pupils to become accustomed to their surroundings. To the criticism that such conditions are unrelated to our public schools, one answer is to show that the interactive events which occur during experimental conditions have the same consequences in a public school setting. The testing of this assertion is a separate demonstration, yet it is an essential research responsibility which cannot be ignored.

A fifth requirement is that the procedures used to assess educational outcomes be closely related to the objectives of the learning activities. That is, when subject matter knowledge is being tested, it should be the same subject matter that was taught. When pupil attitudes are tested, they should be attitudes toward a particular experimental teacher, toward the particular experimental learning activities, and toward the pupil's perception of himself participating in those activities.

A sixth requirement is that the learning activities be scheduled over a period of time which is long enough to obtain between group differences in the measures of educational outcome. In my own research, I have found that 2 weeks of instruction, 1 hour to 1 1/2 hours per day, did provide between-group differences for topics in the field of mathematics, science, and social studies.

Evaluative Teaching Units: The evaluative teaching unit is a label which can apply to a single set of instructional materials, manipulative objects, teacher's manual, and pretests and post-tests that together comprise a logical, coherent unit of instruction. It helps to have different kinds of instructional materials and to have more materials than the teacher can use. This permits the teacher to choose those materials which fit his personal style of teaching, no matter whether it is lecture-assignment, group work, individual seat-work forms, or some combination. The general purpose of such a unit is to control the teaching objectives and resource materials when classes are exposed to different patterns of teaching behavior.

A unit of this type has a great advantage if the substantive content is bizarre, compared with the immediate experiences of the pupils. Thus, we found that a unit of instruction on the geography, history, and economic development of New Zealand was a topic which seldom came to the attention of the Minnesota youngsters who participated in one of our field studies. Given this kind of topic, the odds increased that experiences in other classes, after school, at home, and with TV, did not directly influence what children learned.

It should be recognized that some educators insist that teaching behavior be studied with topics which are a part of the immediate environment of the child. Such a point of view need not be in conflict with evaluative teaching units even though it may conflict with the topic of New Zealand. An evaluative teaching unit could be based on an intensive study of the ecology of one square foot of sod and grass which came from the school grounds or a youngster's backyard. In the deprived section of an inner city, an analysis of the manufacturing processes of products to be found in a local store or the traffic patterns of the transportation system of the city are possible. The undesirable feature of a local topic is that experiences outside of the classroom are more likely to influence post-test performance.

Setting aside the foregoing details, the main purpose of evaluative teaching units is to control factors which would otherwise confound the assessment of teaching behavior. Suppose our purpose was to compare two patterns of using questions and each pattern was assigned to one of two treatments, each treatment having the same instructional purposes and materials. Such a design would surely give our analysis of questions greater precision than would occur if the instructional materials were not controlled.

To suggest another example, suppose there were two groups of beginning teachers,one the product of curriculum A and the other from curriculum B. Let's say that we were interested in knowing whether the college students from these two methods of teacher education perform differently with regard to particular teaching skills. Any reasonable design for this kind of comparison would attempt to control the unit of instruction so that differences in topic and instructional materials would not bias the main thrust of the evaluation.

It is possible to visualize about 20 basic units of instruction covering four contrasting topics and available for five different pupil age ranges from kindergarten through the 12th grade. While this would not adequately represent the wide range of curriculums in the public schools, the units would be extremely useful assessment tools. We could set up normative expectations for particular comparisons. One comparison, for example, could be between instances in which pupils learned more and developed more positive attitudes in contrast with instances in which the reverse was true. These contrasting norms could be established for all types of teachers such as beginning teachers, black teachers with white pupils, inner versus outer city, for different groups of youngsters, and so on. For each meaningful sample, norms could be developed for class formations, the use of time, and any number of qualities which apply to the contacts between the teacher and the pupils. Watching the way that these qualities change during the teaching of a 2-week unit is one way that teaching strategies can be identified.

If evaluative teaching units become standardized assessment tools of teaching performance, they would be relatively expensive to use and quite time consuming. Performance on an evaluative teaching unit could be a useful criterion for discriminating less costly and more efficient assessment procedures. One would hope that certain kinds of microteaching assignments or that the ability to discriminate different patterns of interaction in a sound motion picture would subsequently prove to be highly correlated with performance on an evaluative teaching unit.

The development and use of evaluative teaching units would be an important part of a National Coordinated Program for Research on Teaching Effectiveness. Most of the recorded behavior collected during the first 5 years should be from interactive situations which can be compared. Evaluative teaching units provide the necessary control. Later on, after analytical procedures have been refined, the archive staff can have the luxury of exploring recordings in widely different settings such as open classrooms with self-selected pupil goals and other more radical circumstances.

By way of historical anecdote, members of my research staff have developed three different evaluative instructional units. Two were on the topic of New Zealand, one for the fourth grade and the other for the seventh. A third unit was developed for eighth grade mathematics. Similar control of instructional content has also been used in the form of short 30-minute lessons.[11] One lesson was more like a current events discussion in which the teacher's responsibility was to help pupils express and then summarize their own opinions. The second was more subject matter oriented, like learning to spell specific words or to learn a particular arithmetic procedure; here the teacher was expected to take a more directive and active role.

Increasing the Range of Teaching Behavior

An important research function which any national program would have to provide is the assurance that a wide range of teaching behavior was being investigated.

Current Practice: How teachers act while teaching is a product of their past experience interacting with the present circumstances. Their past experience includes what happened when they were students, what happened in their families, what happened during their teacher education, and their past few years as professional teachers. One can argue that the impact of past experience for most teachers is toward maintaining the status quo, not toward innovation. Similarly, the impact of the teacher's immediate circumstances such as administrative policy, the arrangements of space, the size of the group, the length of a period, the available instructional materials, and the lack of time to think and plan, these forces also support the status quo. In fact, almost by definition, innovation is the product of an unusual experience caused by an infrequent constellation of forces.

As long as research on evaluating teaching effectiveness is restricted to studying representative samples of teachers in our public schools, the more valid conclusions will only be inferences about the status quo. However painful it may be, an objective researcher must at least entertain the hypothesis that current practices may be a poor model of effective teaching behavior and that no matter how thoroughly we study current practices, we will only be reporting the current derangements with ever increasing accuracy.

One conjecture, at this point, is that the greatest contributions to understanding how a teacher can make a difference may result from studies of teaching which are most unlike the public school conditions.

Training Teachers: To me it seems almost self-evident that increasing the range of teaching behavior is a training problem. If we want teachers to act with greater variation while teaching, we will have to train them to do so. The training should have at least two kinds of corrective feedback. First, video/sound recordings should be studied by the teacher for the purpose of refining behavior patterns which are to be performed. Second, pupils should react to indicate whether they thought the teacher was "putting on a show" or whether his behavior seemed authentic.

In a given study, one way to increase the different kinds of teaching patterns is to teach each teacher to act with greater variation during a 1-hour period. A second way is to have teachers teach in their natural style, whatever that may be, but to select teachers so there are significant differences between them. Both alternatives should be explored since the results would have implications for our understanding the teaching strategies and teacher education.

Soar[12] has described curvilinear relationships between a particular educational outcome and the frequency of some teaching behavior. He found a positive linear relationship between a measure of comprehension and a feature of teaching behavior like indirectness at lower levels of indirectness. However, at higher levels of indirectness, the slope of the curve leveled off and finally became negative. This suggests that there is an optimum level of indirectness for an educational outcome such as comprehension. A true experiment to check on Soar's findings would require that teachers be trained to establish a particular level of indirectness or to vary it according to a plan for the duration of a teaching unit. Either another teacher, or the same one with a different group of pupils, would then have to establish and maintain a different level or carry out an alternate plan. In this manner, the curvilinear relationships suggested by Soar could be verified or questioned.

Example of a Pure Experiment: Graham Nuthall[13] and his associates at the University of Canterbury have demonstrated that the planned variation of instructional behavior is a feasible model for studying the associations between interaction variables and measures of learning. Many aspects of the Nuthall experiments can serve as a model for the proposals described in this paper including the magnetic recording behavior for later analysis, planned variation of teaching behavior, and close relationships between the test of learning and the objectives of teaching.

There are older examples of true experiments including Filson,[14] Amidon and Flanders,[15] and Schantz.[16] In each case teachers were trained to perform particular interaction patterns while teaching a group of pupils. The incidence of experiments, compared with field studies, in the research field of teaching effectiveness is shockingly low.

Evaluating Educational Outcomes

The consequences of learning can be found in how the learner thinks, acts, and feels. The consequences of teaching are determined by measuring changes in the thinking, acting, and feelings of students, but these measurements should be carefully restricted by the objectives of instruction. The consequences of teaching and learning can never be completely determined and attempts to measure them are merely estimates based on partial information. When the purpose is to evaluate teaching behavior, the most reasonable assessment of teaching and learning involves the comparison of identical measurements in two contrasting treatment conditions. Here the limitations of our testing procedures which measure only a small proportion of all behavior are reflected in both treatments. There is no attempt to measure what might be called "absolute" learning. The data are collected in order to make the best possible guess about whether more was learned in one treatment compared with the other.

General Rationale: There are a number of interrelated steps in the logic of using evidence to compare the educational outcomes from two or more teaching treatments. Consider the following:

1. We have created two or more opportunities for learning that

have the same objectives of instruction, the same instructional materials, and the same schedule of teaching-learning sessions.

2. We have made careful plans to create differences among these learning opportunities by planning, practicing, and verifying differences in the behavior of the teacher(s).

3. The differences in teaching behavior create differences in the interaction which occurs when teachers and pupils contact each other.

4. We then hypothesize that these different qualities of interaction:

   a. may or may not be perceived by pupils as they occur,

   b. can be perceived by a trained observer or by an analysis of the interaction using a standardized coding system,

   c. will influence the behavior patterns of pupils while they are learning,

   d. will influence the attitudes of the pupils toward the teacher, toward each other, and toward the learning activities,

   e. will alter the pupils' perception of themselves as learners in that particular situation,

   f. will cause differences in performance on the assessment procedures that we use to evaluate the teaching objectives.

5. We believe that observations, interviews, the responses of pupils to test items, and other data-gathering procedures, will adequately sample essential elements *within each item* "a" through "f" in "4" above.

6. We anticipate that an analysis of the data will permit us to make valid comparisons among the two or more opportunities along the lines indicated.

Most readers will note that one obvious omission is any concern for the long range consequences of differences in learning. In some projects, one can repeat assessment procedures in order to obtain measures of recall, forgetting, and persistence of post-test differences. However, many attempts to obtain such measures fly in the face of uncontrolled experiences which undoubtedly influence delayed assessment scores. Such assessments can be recommended when it can be shown that the experiences of the learners between post-test and recall tests are not significantly different for the groups being compared, especially in relation to the teaching objectives.

Even though research on teaching effectiveness has made remarkable progress during the past decade, to my knowledge there is no instance so far in which all the steps in the above sequence have been verified. Yet if we are to face the problems of research on teaching effectiveness squarely, anything less than verifying each step in the logic of evidence leaves room for doubt.

Evaluative Teaching Units: One reason that evaluative teaching units have a strong appeal is that they can include procedures for assessing each step in the chain of evidence, providing the teaching periods are video/sound recorded. In a properly constructed evaluative teaching unit, daily tests of pupil perceptions, observation check lists, pupil attitude inventories, self-concept rating scales, and the more common pretests and post-tests can be designed and field tested along with other parts of the unit. In effect, the unit is custom built for assessment. It is developed and designed to be used for the evaluation of teaching behavior. It can become the foundation for developing more efficient, shorter, and less costly procedures for evaluating teaching behavior.

Random Assignment of Pupils to Treatments: Most of our statistical procedures which have the capability of making adjustments for initial ability, like analysis of covariance and multiple regression, were not designed to adjust for large between-group differences in initial ability. These procedures provide a more sensitive test when the pretest scores are more reasonably distributed. Within reasonable limits, a regression analysis can take into account the variance on the post-test that is attributable to the pretest. Even when successful, however, these adjustments simply fail to control for the quality of interaction, a topic which we will now consider.

The evaluation of teaching behavior is often concerned with associations between variables based on teacher-pupil interaction and some measure of an educational outcome. The discussion in the previous paragraph was concerned with the latter variables. The more common methods of using regression analysis do not effect the interaction variables. Yet it is highly probable that the presence or absence of certain pupil traits will influence teaching behavior. For example, I would guess that bright youngsters will take more initiative, especially with academic problems, and this will create more opportunities for the teacher to respond to pupil initiative compared with a group of slower pupils. The regression analysis of outcome variables does not provide the necessary control of pupil ability when interaction variables are considered. All of these observations argue for the assignment of pupils to experimental sessions in such a way that the groups to be compared differ in initial ability only by small chance occurrences.

Priorities Among Outcome Criteria: The general thrust of performance contracting in education—reviews of research like those of Rosenshine, and verbal statements which I believe I have heard Gage and others make—have tended to place emphasis on the outcome variable of achievement. The motivation is understandable, namely, if we are going to report on what a teacher does that is effective, the criterion by which we judge effectiveness should be learning, because this is the most important outcome.

Achievement, as we measure it under the best of conditions, may not be the most important variable. Aspects of education which may be more important are the desire of pupils to want to stay in school, to have confidence in the teacher and respect for his skill, to see one's self as responsible in learning situations. If there has been a priority placed on achievement between 1968 and 1971, it is surely a case of premature closure.

The goal of this proposal is to insist on collecting evidence regarding educational outcomes at each step in the logical argument by which we infer that differences in teaching behavior will affect achievement. Those steps were listed at the beginning of this section. The implication of this recommendation is that we have no business using measures of achievement as a criterion of effectiveness unless the data at all the earlier steps justify its use. Overemphasizing the importance of achievement is, in my opinion, likely to lead to the identification of teaching behaviors which are associated with bright pupils and not necessarily with effective teaching. For example, when a nationally standardized test of academic ability is used as a criterion measure in research on teaching behavior, we tend to isolate those teaching behaviors which can be found in classrooms which have pupils who are above average in academic ability. Any explanation which suggests that these behaviors are characteristic of more effective teachers should require additional supporting evidence to fill in the logical argument. How do particular behaviors facilitate more learning? What responses are pupils making to these behaviors? Are the attitudes of the pupils influenced by these behaviors? These and similar questions need to be answered.

The Readiness of Our Times for a National Effort in the Field of Research on Teaching Effectiveness

This section considers the readiness of the Nation for an all out effort to determine how a teacher makes a difference. I will first discuss the current state of technological research and development in education which our generation can bring to the task. Next, there are certain political trends which auger well for the proposed program. Finally, there are social and cultural priorities which should be taken into account. Given our technical, political, and sociological readiness, the times appear propitious for a concerted effort.

Technological Readiness

There are three kinds of technological readiness which seem to be directly related to the proposals that are outlined in this paper. First, we have made considerable progress in tools for collecting data about teaching effectiveness. Second, we have increasing information about how the teacher's role is changing and how we can help him learn particular teaching skills. And third, we have more experience with federally directed research programs, how they can be supervised, and how they can be supported. Each of these areas will be discussed briefly.

Research on Teaching Effectiveness

Most of the empirical evidence concerning teaching effectiveness has been secured within the past 10 years. Below are some of the main features of this progress.

1. *Coding teaching behavior*: Systems for encoding verbal and nonverbal behavior in classrooms, training groups, therapy groups, teacher-student diads, and microteaching settings have been published and are now available. *Mirrors for Behavior*[17] and the *Handbook for Research on Teaching*[18] are two references that list different systems. The technology has reached a level at which a display of interaction data lasting up to an hour which can be decoded in only a few minutes, and which highlights particular comparisons that are of interest, can now be provided to a teacher at the instant that the teaching stops.[19]

2. *Coding student behavior*: Modern category systems can provide as many categories as needed for either the teacher or the students. New methods of displaying the data can be arranged to highlight student interaction on forms that are used for the direct recording behavioral events.[20] Systems for tracing the individual participation of a single student have been in use for several years.[21]

3. *Computer applications to coding and decoding*: Direct encoding using pushbuttons which are connected to computers has been reported by Flanders[22] and Smidchens and Roth.[23] An interesting application of computer technology to decoding has been proposed by Collect and Semmel.[24] These contributions have solved the first-round problems with regard to interface and software.

4. *Problems in statistical analysis*: Among the reports of research on teaching effectiveness are several examples of progress in statistical analysis. Soar[25] has shown that some relationships between interaction variables and educational outcomes are curvilinear. Some weaknesses of multiple correlation when applied to interaction data have been identified by Flanders.[26] Computer

technology has greatly reduced problems of data reduction. When data analysis and direct encoding procedures are both accomplished by computer programs, an extremely efficient and powerful data reduction capability is created.

Research on Teacher Education

Along with the foregoing accomplishments in research on teaching effectiveness, we have also made progress in helping teachers change their teaching behavior and in helping college students learn particular teaching skills. This progress in modifying teaching behavior has direct application to the problem of increasing the range of teaching behavior in the programs that are being proposed.

1. *Minicourses and microteaching*: We now have self-directing, individually paced, instructional packages which adults can use to learn particular teaching skills.[27] The progress here helps to ensure that we can create patterns of teaching behavior without being limited to current teaching practices.

2. *Observation in teacher education*: There have now been a series of studies which indicate that teaching behavior is modified by obtaining feedback using systems of interaction analysis. Some of these studies are reported in Amidon and Hough,[28] in Flanders and Simon,[29] and in various professional journals.

The National Scene

There are two trends on the national scene which are directly related to what is being proposed in this paper. One has to do with knowing more about how the role of the teacher will change if we can implement innovation in educational technology. The other has to do with the administration of research programs through the Office of Education.

1. *The changing role of the teacher*: In the development of IPI and CAI at the University of Pittsburgh Center, Lindval and Bolin[30] conclude that—

> "In summary, experience with Individually Prescribed Instruction has indicated that the teacher role in supplementing this system *is absolutely essential** for an individualized instructional program; . . ." (p. 40) (*emphasis not in the original)

The specific skills that are most likely to be needed were described as follows by Cooley and Glaser.[31]

> "The teachers' three main functions have been writing prescriptions for courses of instruction, diagnosing student dif-

> ficulties, and tutoring individuals and small groups of students." (p. 579)

One can speculate that as automated instruction procedures become increasingly available, the skills of responding to pupils, asking questions, clarifying and summarizing, and similar patterns will be used by teachers more frequently. The patterns now most frequently seen in today's classrooms will be decreased and the patterns now least frequently used will be in greater demand. Such trends have direct implications for the programs proposed in this paper.

2. *Office of Education directed research programs*: In his extensive report on the national research effort in education, Gideonse[32] outlines the trend toward research programs directed by the Office of Education (see Chapter VI). With the National Institute of Education standing in the legislative wings (undergoing hearings), and the current and continuing reorganization of the Office of Education, we can expect more active leadership, supervision, and administrative help at the Federal level. This proposal outlines a national program which will require Federal assistance and thoughtful leadership.

Political Readiness

One theme frequently heard when politicians speak on education is evaluation. One message is that we should postpone the passing of new legislation and we should even question expending monies now allocated until we have more evidence that these expenditures will accomplish their stated purposes. One place to start giving examples is with the President.

President Nixon sent a *Special Message on Education* to Congress on March 3, 1970 and in it[33] he said—

> We must stop pretending that we understand the mystery of the learning process or that we are significantly applying science and technology to the techniques of teaching—when we spend less than one-half of one percent of our educational budget on research, compared with 5 percent of our health budget and 10 percent of defense.
>
> When educators, school boards, and government officials alike admit that we have a great deal to learn about the way we teach, . . .
>
> (We should) begin the responsible, open measurement of how well the educational process is working.
>
> To achieve . . . fundamental reform, it will be necessary to develop broader and more sensitive measurements of learning than we now have.

The main theme emphasizes looking at what we are doing, to evaluate our present programs, and to seek more definitive knowledge before we increase our present programs or start new ones. I would assume that what the President and his advisers said about education *in general* applies equally well to research on teaching effectiveness and teacher education. The point of the proposal in this paper is that for the first time in our history, I believe that within a 5- to 10-year period we can promise and deliver substantial progress in finding out what a teacher can do that makes a difference. We have the tools, we now need the resources.

Relationships between a government and its researchers in the field of education are taking constructive turns in other countries as well as within the United States. In fact, research programs in Finland and especially in Sweden set an example of close cooperation. Koskenniemi[34] in Helsinki, Finland, accepted a government request to see if predictors of teacher effectiveness could be measured when college students enter teacher training and to see if these predictors can be related to teacher performance once a person enters the profession. The results did show that Koskenniemi's predictors would involve considerable error if used to select college students. This evidence helped the Finnish Government avoid passing laws that would establish unwise selection procedures based on ineffective entrance requirements. Another example of cooperation between government and an educational researcher is the work of Dahllof[35] in Goteborg, Sweden. He was asked to study the effects of different class groupings on individual differences. Dahllof has attempted to explain why some youngsters "under learn" and others "over learn" and has developed a model which helps in suggesting alternative plans for organizing education which are more efficient. With continuing success, the government in Sweden will itself engage in educational research, since it will try out alternative schemes of school organization based on Dahllof's findings.

One important observation can be made of these two research efforts in Finland and Sweden. In both cases the projects were given financial stability and continuing support over a number of years. It would also appear that there was enough time to formulate thoughtful answers to the questions which originally stimulated the research. Another observation is that in each case the government allocates the task of obtaining information to a person who is qualified to conduct the necessary educational research.

Knowing more about how teachers can make a difference is a pressing problem in this country because every youngster who goes to school comes into contact with teachers and this condition is likely to continue for more than 50 years, even after automated instruction is introduced. Given the possibility of a National Institute of Education or by working with existing centers and regional laboratories, I believe we can start and then expand a National Coordinated Program of Research on Teaching Effectiveness.

Social and Cultural Urgency

The word *relevance* is frequently used these days, especially when we start shouting at each other about social issues. The cry for relevance arises from disagreements about the meaning of events and the consistency with which we use words to describe these events. Thus, if one person describes our activities in Vietnam as "creating peace" and another prefers to call these activities "making war," then a third person can decide which point of view is more relevant to "what is *really* going on." When we talk about our public schools teaching democracy but someone suggests that they are demonstrating authoritarianism, it is quite reasonable to ask which point of view is more relevant to "what is *really* going on."

With regard to the contacts between students and teachers, there are some of us who are no longer interested in speculating about these contacts with the rhetoric of philosophy, ethics, and values. Instead, we prefer the rhetoric of analyzing behavioral events. We hope to begin with a reasonably objective description of the small interactive events which occur during these contacts. We assume that agreement about what is taking place is more likely to occur if we restrict ourselves to behavioral description. These descriptors may then lead us to meanings, and thence to labels, and finally to generalizations. The hope is that generalizations identified by this more inductive process, that is—starting with the small behavioral events and arriving at generalizations, that this kind of thinking will help us to create a point of view which is more relevant. It will more accurately reflect what is *really* going on.

Most of us who have studied and analyzed the interactive events which occur when students contact teachers have been deeply disturbed by what we have discovered. Teachers tend not to listen very well when students speak. Teachers respond to ideas expressed by students less than 10 percent of the time. Most of the time the only way a student can interact is by limiting his participation to responding, since there are very few opportunities to initiate. The superior-subordinate aspects of teacher-pupil contacts are probably more pervasive than the content being learned and, I believe, have a more profound impact on our youth than any other feature of schooling.

To whatever extent there is a discrepancy between what *really* goes on during teacher-pupil contacts and how we educators choose to think, then our thinking is not relevant. If our thinking about the current interaction in our classroom is not relevant, then it would seem to me that our chances of making substantial improvements are poor indeed. To avoid this sad state of affairs, we need to know the details of teacher-pupil interaction as it currently exists and we also have to investigate new patterns which we hope will bring about improvements.

The fundamental questions which are raised by this paper can be simply stated: are we sufficiently concerned about the present state of affairs so that we will find the resources necessary to mount a national effort? Are the proposals to be found in this paper of sufficient merit so that they can be incorporated in a national plan?

Summary

A plan is presented in this paper to create a National Coordinated Program for Research on Teaching Effectiveness which will be run by the staff of a National Cooperative Archive for Recorded Teaching and Learning Behavior. This proposal is based on the assumption that we do not now know how to answer the question, "HOW do teachers make a difference?" Even though we have made more progress in this area of research during the past 10 years than we have made in the history of mankind, this progress can best be summarized by saying we now know how to proceed in order to discover HOW teachers make a difference. We have new tools for research as a result of work during the past decade. We have good hunches about what to look for as a result of the work during the past 10 years. We know more effective ways of helping a person develop teaching skills as a result of the past 10 years. We now know that definitive research will require greater experimental control than we have exercised during the past 10 years, especially with regard to the academic ability of children.

During the 1970s we can push farther toward answering the question, "How do teachers make a difference?" However, I believe that this will take a national effort and cannot be left to individual researchers, or, for that matter, to one or two centers.

References

1. Kaplan, A. *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., 1964. pp. 71.
2. *Ibid.*, pp. 53.
3. Thelen, H. A. *Education and the Human Quest*. New York: Harper Brothers, 1960. (See Chapter 3, pp. 34.)
4. Brodbeck, M. "Logic and scientific method in research on teaching," in N. L. Gage (Ed.), *Handbook for Research on Teaching*, pp. 48.
5. Simon, A. and Boyer, E. G. *Mirrors for Behavior*. Philadelphia, Pa.: Research for Better Schools, 1967. (XIV volumes + Summary)
6. Flanders, N. A. (*et al.*) "Teacher influence patterns and pupil achievement in the second, fourth, and sixth grade levels," Vol. I. Terminal contract report, School of Education, University of Michigan, Ann Arbor, 1969, pp. 10.
7. Smidchens, U. and Roth, R. "Michigan-Ohio regional educational laboratory's computerization of interaction analysis project," Mimeographed by MOREL, 3750 Woodward Ave., Detroit, Mich. Now out of print.
8. Massialas, B. G. (*et al.*) "Structure and process of inquiry into social issues in secondary schools," Vol. III Terminal contract report, School of Education, University of Michigan, Ann Arobr, Mich., 1970, pp. 66ff.
9. Borg, W. R., Kelley, M., Langer, P., and Gall, M. *The Minicourse: A Microteaching Approach to Teacher Education*. Beverly Hills, California: Macmillan Educational Services, Inc., 1970.
10. Flanders, (*et al.*), *op. cit.*, pp. 100ff. (See Ref. #6)
11. Finske, Sister M. J. "The effect of feedback through interaction analysis on the development of flexibility in student teachers," unpublished doctoral dissertation, University of Michigan, 1967.
12. Soar, R. S. "Optimum teacher-pupil interaction for pupil growth," *Educational Leadership*, 1968, *26*, pp. 275-280.
13. Nuthall, G. "Experimental studies of teacher behavior," mimeographed draft supplied by author, Department of Education, University of Canterbury, Christchurch, New Zealand, 1971.
14. Filson, T. N., "Factors influencing the level of dependence in the classroom." Unpublished doctoral dissertation, University of Minnesota, 1957.
15. Amidon, E. J., and Flanders, N. A., "The effects of direct and indirect teacher influence on dependent-prone students learning geometry." *Journal of Educational Psychology*, 1961, *52*, pp. 286-291.

16. Schantz,Betty M. B., "An experimental study comparing the effects of verbal recall by children in direct and indirect teaching methods as a tool of measurement." Unpublished doctoral dissertation, Pennsylvania State University, 1963.
17. Simon, A. and Boyer, E. G., *op cit.*
18. Gage, N. L. *Handbook for Research on Teaching*. Chicago, Ill. Rand McNally and Company, 1963.
19. Flanders, N. A. *Analyzing Teaching Behavior.* Reading, Mass.: Addison-Wesley Publishing Company, 1970.
20. *Ibid.*
21. An interesting recent example is a study by the Civil Rights Commission, Henry Ramirez, Director, in which the teacher-pupil contacts with Mexican-Americans, blacks, and Anglos was recorded separately.
22. Flanders, (*et al.*), *op. cit.*
23. Smidchens, U. and Roth, R., *op. cit.*
24. Collet, L. S. and Semmel, M. I. "The analysis of sequential behavior in classrooms and social environments: problems and proposed solutions," mimeographed prepublication draft supplied by authors, 1970.
25. Soar, *op. cit.*
26. Flanders, (*et al.*), *op. cit.* (See Ref. #6)
27. Borg, (*et al.*), *op. cit.*
28. Amidon, E. J. and Hough, J. B. *Interaction Analysis: Theory, Research, and Practice*. Reading, Mass.: The Addison-Wesley Publishing Company.
29. Flanders, N. A., and Simon, A. "Teaching effectiveness: A review of research 1960-66." In R. L. Ebel (Ed.), *Encyclopedia of Educational Research.* Chicago: Rand McNally, 1970.
30. Lindvall, C. M. and Bolvin, J. O. "The role of the teacher in individually prescribed instruction," *Educational Technology*, 1970, *10*, pp. 37-41.
31. Cooley, W. W. and Glaser, R. "The computer and individualized instruction," *Science*, 31 October 1969, *166*, pp. 574-582.
32. *Educational Research and Development in the United States.* U.S. Department of Health, Education, and Welfare, Office of Education. Superintendent of Documents Catalog No. HE5. 212:12049, $2.00. Washington, 1970.
33. All quotations are from *The New York Times*, March 4, 1970, p. 28ff.
34. Koskenniemi, Matti. *The Development of Young Elementary School Teachers: A Follow-up Study*. Suomalaisen Teideakatemian Toimituksia Annales Academiae Scientiarum Fennicae, Sarja-ser. B Nide-tom. 138. Helsinki, 1965. 635 p.

35. Dahllof, U. S. *Ability grouping, content validity, and curriculum process analysis*, Project Compass 13, the Institute of Education, University of Goteborg, No. 7, June 1969.

Chapter 7

LEARNING ENVIRONMENTS-OR-ROOMS FOR THOUGHT

Lawrence M. Stolurow

There are many answers to the question "How do teachers make a difference?" We have heard several from the preceding speakers in this conference. What follows here will hopefully add and possibly integrate some of the points already made.

Interventive and Transactional Instruction

My belief is that there are two rather different points of view that underpin the relationship between the student and the source of information that provides educational experiences. One of these conceptions is that teaching is a process of intervention. This conception seems to be derived from the view that the basic process of development both physically and intellectually takes place naturally as the individual grows and matures. Instruction is seen as an intervention in relation to the individual's natural development. Although he does not use the term, Bruner (1964) describes one form of intervention as follows: "Instruction consists of leading the learner through a sequence of statements and restatements of a problem or body of knowledge that increases the learner's ability to grasp, transform, and transfer what he is learning" (p. 313).

The second conception is the one presented here. In some ways it is newer, and in many it is qualitatively different from *interventive instruction*. This alternative conception might be called *transactional instruction*. Teaching from this point of view is a process of interaction, of give and take, to achieve some end, but the interchange is not necessarily between a student and a teacher. Although student-teacher interactions may be one of the more important forms in which transactions occur at the present time, other forms appear to be potentially more important. Teaching in transactional instruction involves the management of the environment, both social and physical, to provide maximal opportunities for various kinds of self initiated transactions to take place between the learner and materials from which he is learning.

Psychological Factors in Transactional Instruction

Transactional instruction is primarily psychologically oriented and focuses on the processes used by the learner and the catalytic role of the instructor. In other words teaching is to facilitate learning but not to become a part of it. Just as there are many kinds of chemical catalysts there are many kinds of instructional catalysts. Each of them has a different form. All of them are alike in that the instruction is specifiable and repeatable and it facilitates learning and understanding. For transactional instruction processes are reduced to either a heuristic or an algorithmic form. The common purpose is to provide learners with opportunities to learn under conditions that maximize initiative on the part of the learner and minimize the intrusion of both the teacher and the means of instruction.

Engineering of Transactional Programs

In transactional instruction the teacher is more of an engineer or manager than a source of information. He structures and delimits the learner's environment using such devices as he can invent or discover to focus attention, to restrict activity, to provide useful references or resource materials such as experts or critics. He may develop a walking plan to provide specific types of experience—laboratory demonstrations or projects or synthetic experiences through films, slides, audiotapes, or ETV. With a well developed set of objectives the teacher utilizes his ingenuity and most appropriate resources to develop a set of experiences appropriate for each child. The rule of thumb followed is to maximize the probability that the student learns what is contained in the objectives and to minimize the amount he is told about something or required to memorize for memorization sake. An implicit assumption is that the student's response to a problem is qualitatively different, particularly in the cues that initiate it, and in the organization of his thinking, from his response to information he is told to learn or to use when he solves problems. A problem defines a context, a set of conditions and delimits the set of relevant and useful cognitive activities. Information and activities automatically become meaningful to the student. The teacher, in developing the problems, has to analyze experiences to determine their implications for what the student is thinking and for the information required. The teacher is engaged in more planning of both real and symbolic experiences than when obliged to develop a lesson plan for a lecture or oral presentation to a class. The teacher analyzes problems so as to anticipate both correct and incorrect approaches and as a result becomes involved in a deeper conceptualization of student learning problems and also becomes better able to interpret student errors when they occur. The teacher writes more scripts than outlines and is more laboratory than lecture-oriented. He also becomes more diagnostic and prescriptive and therefore more oriented toward individual students than when he "teaches a class."

Materials designed for use in transactional instruction are planned with either as much, or more, concern for what students do while learning as for the accuracy and completeness of the content they are given. Furthermore the transactions are interactive and not just responsive. This means they not only conclude each interaction but also provide for the next one. In transactional instruction rules, or principles, are used to systematize learning on a step-by-step basis. However, this is accomplished as a structure with many paths to each goal. This is accomplished by providing the learner with many options at every choice point except when it has been determined that a single path is the best. At the beginning, in the design stage, the question is "What are the potentially most effective combinations of content and process for the students who will be given this program?" In this way the presentation or instructional strategy questions are faced from the start and the materials that best support the important considerations relating to teaching are developed. These questions have important implications for the information processing activities of the students. If the programing problems are not faced from the beginning or at all, then a large number of problems are produced, but one of the more significant ones of these is that the information will be presented as many ways as there are differences among the teachers who are allowed to use it on a provisional basis. The blending of principles to make up strategies has important implications for the development and use of psychological constructs in instruction and for the development of instructional materials.

Computer programs for CAI systems make it necessary to deal explicitly, and in operational terms, with both the matching and interrelating of pedagogical principles with specific segments of content. The process begins with the development of the objectives to be achieved by a student. In developing CAI programs for transactional instruction, the requirement is to be explicit about both the psychological and the epistemological problems of instruction. Materials designed for transactional instruction require sophisticated engineering of both the information and the processing. One compensation for the extra effort required to design these materials and build them is the reduction in the risk of failure to attain the objectives. Also built into the procedures for their development is the formative evaluation process that provides for instructional accountability. Not only is the material organized, but also the requirements for implementation are explicated. Often more than one form of presentation is developed to provide for differences among individual learners. An additional developmental burden over the requirements of normal curricular development is the design and development of remediational instruction. Not only are normal learning problems anticipated and provided for, but also learning disabilities. Thus the curriculum development represents a whole set of developmental activities as they are partitioned by current practices. By considering them as a set there is greater assurance that something is being done about the total range of possible problems than if the developmental planning were only for the "normal" or "average" learner.

Materials designed for transactional instruction have to be conceived for use in a prescriptive mode. They have to be validated for the variety of learners for which they will be used as well as for the subject matter covered. Both epistemological and logical considerations have to be taken into account in the frame of reference of an interaction or dialogue.

Materials designed for interventive instruction are complete once the content is specified and an outline has been developed for their presentation. Typically a scope and sequence chart is prepared. Materials used for transactional instruction are characterized and developed in greater depth. Units that are not only coherent within themselves but also as sets have to be designed and developed. Typically these units have to be conceptualized as elements in a variety of sets each of which provides an instructional interaction with a student possessing certain characteristics at the time the decision was made to engage him in one transactional script rather than another. The various possible sequences provided for have a rationale. They are achieved through the use of contingency rules for handling response-produced information (both feedback containing knowledge about results achieved and new information). Criteria are included for use in evaluation as well as for use in making on-line decisions about branches. Both criteria and contingency rules are "built into" the materials. In a sense the contingency rules define the structure.

Interactions provided by a teacher or a computing system as a game or simulation are opportunities for thinking. The overt behavior is really not the critical aspect of student response. More important is the conceptual infrastructure that involves symbolic activity and information processing skills.

Control and Responsibility for Instruction

It is important to distinguish between two rather different conditions of control and responsibility for the instructional experience. Whereas interventions are determined by the teacher or the author of the instructional materials, in transactional instruction control and responsibility is shifted from the teacher or teaching system to the learner. An important part of what the student is expected to learn is control and responsibility for his own instruction.

By encouraging the transfer of control and responsibility from the system to the student the learning experience involves initiative and in that sense it has important affective objectives as well as cognitive ones. The author of a transactional instructional program is clearly responsible for the objectives and for the delimitation as well as the characterization of the infrastructure of the instructional system. However, the particular way in which the materials unfold for a student are increasingly his responsibility as he progresses through a program.

Objects, Symbols, and Devices for Instruction

Learning environments are more heterogeneous than they are homogeneous in transactional instruction. A variety of objects, symbols, and devices are used and frequently there are alternatives for each of these kinds of materials so that preferences can be honored. Books, laboratory specimens, and physical and chemical substances are static materials. Films, videotapes, games, simulations, and role-playing are dynamic materials. Transactional instruction relies more and more upon dynamic materials. If a computer-based instructional system is used to provide instruction it can contribute fluidity and momentum to experiences. It also can provide the record keeping.

In writing about computer-aided instruction (CAI) one author described a type of transactional instruction as follows: "In the CAI context a teaching system is a broad concept indeed. It includes student interaction with a variety of media—audiotapes and photographic images as well as text—and the use of different modes of instruction at different times in the teaching process, depending upon a student's performance. Unlike a film, ETV or lecture, CAI is a response-dependent teaching system." (Stolurow, 1969, p. 66)

Not only is the environment created for the learner by a CAI system heterogeneous, but it is typically used interchangably with a teacher and media. The sum total of the mixture is a much richer sensory experience and therefore a qualitatively different internal representation of phenomena than is provided by a homogeneous and less dynamic set of experiences.

Teachers and Transactional Instruction

Learning environments have to be created for teachers as well as for students. Teachers need to solve problems like those which they will give to their students and they need to also solve problems involving the design of problems for particular kinds of students. The problems for teachers have to be developed carefully since they will serve as models for the teachers to use when they create their environments for students.

It is probably accurate to say that with conventional approaches to teaching teachers are taught to intervene. They are taught to "tell," "lead," and "direct" students. They are not taught the essential skills for transactional instruction. For example they are not taught to identify and formulate interesting and useful problems. Problems have to be expressed in ways that are meaningful and yet interesting to the learner. A student of a given ability level must be able to cope with a problem and must be able to feel a reasonable degree of confidence in dealing with the problem.

Some guidance is provided teachers by research. If a teacher wants to use the transactional approach to instruction some guidance is available. Since not all of the problems to be developed will be cognitive and many which are will have affective components, it is important to have a familiarity with the conventional wisdom with respect to these domains (e.g., Bloom, et al., 1956, for the cognitive, and Krathwohl, et al, 1964, for the affective domain). In addition it is useful to have knowledge about the ways in which affective factors influence learning if instructional programs are to be developed to meet individual needs. For example, the socioeconomic background of learners is known to affect their motivation level for school work, particularly their achievement motivation (Rosen, 1959; and Strodtbeck, 1958). The poorer motivation of students from lower socioeconomic (occupational) groups suggests that their teachers need to be aware of this problem so they can deal with it. Some guidance could be provided by Shaftel, Crabtree, and Rushworth (1960). In their opinion the teacher must do these four things: (1) make sure that the emotional climate of the classroom is suitable for the development of a healthy self-concept; (2) evoke problems that are not immediately apparent to the children; (3) stimulate a problem solving climate, which involves the process of search, rather than focusing on one right answer; and (4) plan a curriculum which stimulates problem solving by the use of experience units, construction activities, science experiments, group work, dramatic play, and role playing. Obviously the components of this mixed strategy need to be developed separately and validated. However, while research is needed to bolster these findings, they corroborate the general conception of transactional instruction to provide interim, if not more enduring guidance.

The exploration of alternative solutions to problems appears to be a very basic mechanism (e.g., Stolurow, 1946). The desire to learn and the interest and willingness to undertake problem solving appears to be influenced, if not based, upon cultural, motivational, and individual factors. It would appear that instruction provided by appropriately designed learning environments could facilitate the acquisition of strategies useful in the exploratory phase of problem solving, and might, for example, through repeated opportunities to solve particular kinds of problems through gaming and simulation, lead to the development of useful cue priorities in that their use would, on the average, improve the individual's rate of problem solution. While the predisposition to explore alternatives, which is a precursor to the development of a cue priority, is probably most affected by cultural and motivational factors, it also is true that the base level of a student probably can be elevated by the way in which his learning environment is constructed and interacts with him. Curiosity is a response to uncertainty and to ambiguity (Berlyn, 1960). If a task is something the individual has done frequently, there is little exploration; if it is new and too ambiguous or uncertain, the effect may be to elicit anxiety and to thereby promote confusion which would impede problem solution. In order for the individual to willingly explore a problem solving situation there has to be an optimum level of uncertainty about it. Once the process is started, the activity must be maintained until mastery is achieved. Some processes require careful management, judicious use of reinforcement schedules, and reinforcers, and the maintenance of a tolerable level of ambiguity. To keep the learner's level of activity up and to promote learning-how-to-learn, motivation should be kept at a near optimum level. The maintenance of sufficient motivation to produce exploratory activity depends upon the ratio of benefits from response to the perceived degree of risk in making a response or commitment to a plan. From this point of view, one role of the teacher is to promote the reduction of uncertainty by such procedures as providing useful information. The learner derives encouragement from the perception that the task involves low risk. Consequently, a teacher could contribute to the learning process by reducing the fear of failure. Many options exist for the teacher to reduce the fear of failure. He can develop a good personal relationship with the learner; he can group learners in ways that provide comfort to the anxious learner. He can make the consequences of error appear to be less personal so they do not imply personal intellectual failure or character defect. Negative implications as feedback from attempted solutions could result in reduced exploration of alternatives in future problem solving situations.

In transactional instruction, the teacher and/or the system provide the student with a sense of direction by giving him feedback that reveals the relevance of information obtained from response, and this can be done without actually providing the student with the right answer as many systems do. One model comes from the child's game of "hide-and-seek." In that game the individual is told he is "hot" or "cold" when he responds. This "feedback" reveals the relevance of the response without giving away the answer. Thus it serves to support exploration during problem solution. This type of feedback is useful to the learner in orienting him to the solution. There are many ways in which learning environments can provide orienting information to students during problem solving. The preferred methods are those which the learner can internalize and then use himself.

Another role of the teacher in transactional instruction is to build into the learning environment a variety of useful reinforcing mechanisms. One way in which this can be accomplished is by arranging for alternative presentations of problems which are simpler and arranging for their use whenever errors occur from attempts to solve the more complex form. By either accentuating or developing the perception of the difficulty level of a problem, e.g., telling high school students that the problem is a difficult one for college students to solve, the effect of solution can be accentuated. The effect of correct response is greater when the problem is perceived by the learner to be a difficult one. It also is possible to actually reward the solution of difficult problems in a more emphatic way. The degree of reinforcement can be built into a problem-solving task by adjusting the discrepancy between the student's ability to solve a type of problem and its complexity level. The greater the discrepancy the greater the amount of uncertainty, in general, and the greater the reinforcement from correct response.

The problems that are most difficult for a learner have the greatest uncertainty. If the uncertainty level gets high then anxiety and fear of failure increases. If the problem is complex then increased anxiety tends to reduce the probability of solution. There is a reciprocal relationship between the anxiety level of the learner and the likelihood of solving a complex problem. On the other hand, if the problem is relatively simple for the learner then increased anxiety would promote learning. By being present and available for certain kinds of help the teacher can reduce the student's anxiety and thereby increase the probability that the student will solve the problem if it is complex. The availability of a knowledgeable peer, or even an interactive computer system could serve to reduce the learner's anxiety level. Building into the system ways of detecting relatively complex problems and of identifying errors that suggest a change in strategy to provide a reduction in anxiety would make transactional instruction more effective.

Designing Learning Environments

The design of learning environments is an emerging technology. It is not only achieved but also developed through the use of computers that can be programed for interactive instruction. In order to make an interactive instructional program operational, sets of possible contingencies need to be described at a level of detail that is greater than that required by alternative approaches. This process of explication makes the computer uniquely suited to the task of developing concepts that are now latent or only manifest at a gross level and in need of refinement. In order to implement them on a CAI system it is necessary to make the concepts very explicit. Consequently, the testing of their effectiveness is a way of validating the contingency.

There are many steps involved in the process of designing learning environments. A task analysis is a typical first step. From the analysis a set of behavioral objectives is developed. Then test items are prepared to provide a criterion referenced test by means of which both the effectiveness of the instruction can be evaluated and whereby the voids in each student's learning can be determined. A hierarchy is developed to represent the basic structural form of the objectives and to reveal priorities that should be observed in the development of an instructional program or set of materials for guidance. It often is erroneously assumed that the hierarchy of objectives is a necessary or sufficient specification of the instructional program. The same structure representing one set of objectives can, in fact, generate a large number of instructional sequences. In addition to necessary content and skill priorities the sequences used for instructional purposes must have a number of additional bases.

In developing course material attention has to be given to making it easily learned by different ability levels. Bruner (1964), for example, has identified three factors affecting the ease with which a domain can be learned. These are: "mode of representation," "economy," and "effective power." (p. 309-310). For him there are three modes of representation of a domain of knowledge. They are: (1) a set of actions appropriate for achieving a certain result—"enactive representation," (2) a set of summary images or graphics that stand for a concept without defining it fully—"ikonic representation," and (3) a set of logical propositions drawn from a symbolic system governed by rules or laws for forming and transforming propositions—"symbolic representation." Not all domains of knowledge are equally well represented by these three modes. For example, political science and law are hard to diagram whereas architecture and geography lend themselves to imagery more than to verbal or mathematical analysis and description.

Economy is the amount of information that must be held in mind and processed to achieve comprehension. Formulas are considered more economical than long verbal or tabular representations. Physical presentation may affect the economy of transformation, speed of processing, and transfer of knowledge learned (e.g., Stolurow, 1967). For example, students taught symbolic logic by using the Piano-Russell notation system and those taught the same concepts by using the Polish notation learned at different rates and showed different degrees of transfer of their knowledge to new tasks. These differences were traceable to the differences in transformational economy which the two different symbolic systems required of the learner. Also contributing to the economy of learning are "mnemonic tricks" (Miller, 1967) which are a misnomer in that the psychological processes involved in their use are the same as those involved in developing the laws of nature—". . . all complex, symbolic learning proceeds in this way. The material is first organized into parts which, once they cohere, can be replaced by other symbols—abbreviations, initial letter, schematic images, names, or what have you—and eventually the whole scope of the argument is translated into a few symbols which can all be grasped at one time."

Some measure of the extent to which mental economy can contribute to learning a task is indicated by the results reported by Smith and Miller (see Miller, 1967). They found that learners required 20 percent fewer trials to learn strings of binary symbols, than strings of the same length drawn from either a 32-item population or an 8-item population. They also confirmed findings of Brogden and Schmidt for the 32-item and 8-item tasks, namely, that the length of the list which had to be learned was critical and not the number of alternatives offered at each point. The exception is when there are only two alternatives for each choice point, then the learning is easier. This finding it seems is due to the fact that memories are limited by the number of units or symbols that must be mastered, and not by the amount of information that these symbols represent.

Teachers can make a difference if they teach students to: "organize materials before memorizing it so that the same amount of information is represented in as few symbols as possible." This is called "unitizing." The labeling process, or symbolizing, also is a basic problem in learning, and teachers could make a difference if they taught students to unitize and symbolize. Extrapolating from natural languages it would appear to be useful for the student to develop hierarchies of cognitive units. To do so would make their mastery more economical. An instructional theorem might be stated as a paraphase of John Locke: present a small number of things to be learned, connect them together into a single unified idea and then give the set a name. Teachers also could make a difference if they presented only a limited set of terms, ideas, concepts at one time since it is the length, not the variety of the material, that makes the task more or less difficult. In summary, symbolizing and organizing are pervasive intellectual activities and both teachers and instructional programs should be aimed at teaching students ways in which to perform them more efficiently.

"Effective power" (Bruner, 1964, p. 313) is the generative value of a set of learned propositions. Lists presented for rote learning are considered inert in generative power and therefore low in effective power. Models and theories are potentially rich in effective power. Effective power is assessed in terms of the learner's ability to relate apparently unrelated things once he has mastered a concept, model or theory, or a system of analysis or synthesis. Teachers could make a difference by providing students with integrative models and by teaching students to organize material.

The Message of a Medium is its Information Processing Requirements

Important in the design of instructional systems or learning environments is the conception of the media employed in presenting information to the learner. An important difference between the systems designed for instruction, as contrasted with those designed for use in entertainment or business, is that the message of media is the information processing activities produced in the learner. An analysis developed by Tosti and Ball (1969) sets the stage. They made a logical and functional analysis of instructional systems with respect to information processing. From this point of view, the three basic capabilities of a teaching system are: (a) the transmission of instructional information (stimulus capability); (b) the accepting of measurable behavior of the student (response capability); and (c) the changing of presentations based upon the behavior of the student (management capability). Each capability may be further differentiated in terms of two attributes: (1) form and (2) frequency. This analysis results in the basic 3 X 2 matrix represented in table 1. This table also contains a deeper analytical description of some of the potentially interesting variations in presentation. Teachers as well as interactive computer-based systems for instruction should be able to provide the functional capabilities identified in this table. In addition, they should provide the necessary recordkeeping for use in analyzing student response histories. These data can be used for either making "on-line" decisions at the moment or in research designed to improve instructional programing. The following discussion is not limited to the Tosti-Ball analysis of problems; rather, it uses their taxonomy as an infrastructure in presenting a set of psychological problems associated with the design of interactive-instructional systems.

Stimulus form (representation). Most directly related to media is stimulus form or mode of sensory reception by the learner of the instructional material. Once a teacher has developed a form of symbolic representation of concepts and materials, it is necessary to represent it in one or more forms for presentation to the learner.

Stimulus frequency (duration). The duration of presentation, from transient to persistent, is interesting educationally. Movies usually, but not always, are transient presentations whereas printed text is relatively persistent presentation. A classroom lecture is more transient than a "chalk talk." While these distinctions are relatively general they point up a variable that can make a difference in the reaction of a learner to a presentation of materials.

Response demand. The student's responses to stimuli, while not always unambiguous nor sufficient for inferential purposes, are nevertheless what the educator has to use in arriving at decisions about teaching and student learning. Responses of greatest interest educationally can be categorized as overt-written, overt-spoken, and covert. Theories of learning differ in the extent to which they utilize and depend upon response-derived information as a basis for inferences that are made regarding the unobservable mental processes mediating the observed performance of the learner (e.g., see Hilgard and Bower, 1966). Similarly, teachers and instructional systems differ in their use of response data. The demand for responses in terms of form and frequency provides the critical set of conditions for instructional decisions and therefore is an important aspect of design.

Response demand frequency. This dimension describes how frequently the student is expected to respond (either overtly or covertly) in a given period. A variety of schedules have been identified and studied by Skinner (1961) and others (e.g., Staats and Staats, 1963). Many of the effects which variations in frequency have on shaping behavior are well known. Also involved in the predictability of events that follow response. Aperiodic schedules have a strong effect upon the learner's performance. In cognitive learning there seem to be analogues in that questions demanding response as well as opportunities to question presentations improve learning and retention.

Management form. The decision to present a specific learning exercise to a given student based upon an assessment of some behavior of that student is the critical element in instructional management. This process of decisionmaking can be either "instructional system centered" or "student centered." If it is centered in the instructional system the interactive strategy is determined by the teacher or author of a program. This means that decisions regarding the content sampled and the sequence used, as well as media selected, are all predetermined and knowable both in advance and in detail regardless of whether the student has used them or not. System center decisions can be sensitive to student response or insensitive to them. In addition, the management decisions can be repetitive, or not; they can be at one or many different levels of difficulty or complexity; they can take different forms; and they can be diagnostic as well as prescriptive. Each of these is a management concept and it is implemented, or not, in a system by programing specific rules (e.g., Stolurow, 1969). Even system centered strategies can be adaptive (see Stolurow and Davis, 1965). Adaptivity in this context means that the contingencies specified in advance are not particular events but rather sets of possibilities.

Bryan's (1969) "ad lib CAI" is comparable to Stolurow's (1969) "problem-solving mode of CAI." In this mode the student is given access to the computer equipped with one or more languages and possibly a library of routines and he is in control. Environments in which student management exists have been provided by a simple-to-learn, general purpose language such as LOGO (Feurzeig and Papert, 1970) and PILOT (Starkweather, 1969) and also in a more limited substantive context defined by a particular subject matter that an author programed in modular flexible structure using a general purpose CAI language, e.g., Grubb (1969) for statistics and Manwell, Daugherty, Desch, and Stolurow (1971) for Russian.

Systems have been designed that permit the sampling and the sequential decisions to be made by either the system or the student and for this to shift back and forth throughout a program (e.g., Carbonell, 1970). This is a mixed strategy of management.

Management frequency. The relative frequency with which management decisions are made represents a basic factor in differentiating systems. For example, decisions can be made on-line in real time or off-line in machine time. When decisions are made on-line during learning, the decision process typically involves branching and the instructional system is called a CAI system.

However, when decisions are made between learning assignments of rather substantial duration and are accomplished off-line, or even on-line, but not under real-time conditions, the system is called a CMI system, or a computer-managed instructional system. Since CMI is not accomplished in real time, larger units of material separate decisions and therefore there is a wider spacing of these decisions. CMI is "coarse grained" management, while CAI is "fine grained." This is not a dichotomy; the descriptions used represent the extremes between which there are many variations.

Psychological Not Industrial Distinctions are Needed

From this brief analysis, it is apparent that different questions are asked in designing an instructional system from those asked in designing other systems using media. In determining the benefits to be derived from a system one looks for different things from an instructional system. For example, from the psychological point of view a proper question is "Which presentation duration is more effective for teaching X?" where X is a particular subject matter described in terms of behavioral objectives. This question is considered more useful than "What medium is more effective?" The latter is extended to instruction and asked in connection with the teaching of particular subject matters. However, it is not a useful question. Not only are different parts of a subject matter likely to be quite different in their requirements for responses, but also a medium such as film might run for 5 minutes as a single concept loop, or for 50 minutes presenting a large number of concepts and facts. Medium is not a denotative term for instructional purposes. There are many variations possible in the form and use of each medium, any one of which could have quite different psychological implications from the other.

Studies that have attempted to analyze differences between media, rather than in functional terms (like those contained in the Tosti-Ball taxonomy) have been unsuccessful. For example, Mac Lennan and Reid (1967) reported the abstracts of 350 media studies which were mostly television and film comparisions. Almost none reported significant differences due to media. Campeau (1967) reviewed the literature involving comparisons among television, film, conventional lectures, programed instruction, pictorial presentations, radio recordings, three-dimensional models, and field trips. The majority of the media studies reported no differences in student achievement. Where differences were reported they were as often in favor of one of the media as they were of the other except for programed instruction. A number of surveys have reached essentially the same conclusion (e.g., Barrow and Westley, 1959; Halloran, 1964; Holmes, 1960; Kumata, 1956, 1960; Schramm, 1962; Stickell, 1963). The only conclusion that can be drawn from nearly a thousand of these studies is that there is no basis for differentiating media, as media, for the improvement of the effectiveness of instruction.

From a psychological analysis in terms of functions one possible explanation is that the comparisons have been between heterogeneous sets of conditions and what was labeled film, or ETV, for example, in one study was not functionally equivalent to a presentation carrying the same label in another presentation. Also from study to study these labels are crude specifications for heterogeneous areas of instructional content to which they have been applied. Furthermore, a basic methodological problem exists in such comparative studies. This has been pointed out by Stolurow (1964) and Holland (1965). Each study consists of a comparison of only one example of the medium which could hardly represent the population of variations that is referred to as the medium. A film on inertia is only one example of the way film can be used to present this concept. In short, the media studies have dealt with the problem by asking the wrong question. In addition, they have suffered from more widely recognized methodological flaws (e.g., Ellis, 1962). The result, of course, is that the state of the art is not very advanced for instructional purposes and the need still exists for data. Interactive systems and software now permit one to do the kind of research that is needed to answer the real questions. Computer-based systems have been designed to permit an analysis of the variables that are involved in defining the problems psychologically rather than industrially, the way they have in the past. Film, television, and audiotapes are not psychological variables and they are not a sufficient designation of a set of variables for either research or teacher training to insure reliability or replicability from one exemplar to another in a study of instructional effectiveness.

Information processing capacity. A major problem of fundamental importance in studies comparing the effectiveness of auditory, visual, and audiovisual presentations for instructional purposes is whether or not a human being can handle the simultaneous presentation of audio and visual stimuli. There are differences of opinion among psychologists. Broadbent (1958) for example, takes a "single communication channel" conception of human information processing. For him the audio and visual stimuli are analogous to balls arriving at a flap valve that has two channels of input and one of output—a y-shaped system. If the audio and visual arrive simultaneously, jamming will occur. His over-simplified conception fails to take into account amount of information and redundancy. Jamming can occur when the information to be processed exceeds the information processing capacity of the central nervous system. His model can explain the loss of information when information has to queue up to be processed (Osborne, Quastler and Tweedel, 1955; J. G. Miller, 1963).

Garner (1962) does not accept this view and argues for multichannel processing in the human being. It is fairly well established that man has a limited capacity in information processing as determined by both the modality and his central nervous system. While direct measurement is impossible, indirect measurements have been made. Man's information processing capacity appears to be quite stable in the adult and it appears to level off at between two and three bits per sec. (Frick, 1953). This corresponds to "the magic number seven, plus or minus two" (G. A. Miller, 1956). This means that the normal adult's information handling capacity is within the range of five to nine elements or units at a time.

The important question from the instructional systems point of view, however, is not the limitation of man's capacity for information processing, but rather how to utilize this capacity to its fullest extent. If one could keep up a steady rate of two to three bits/sec. during his normal working life he would store a phenomenal amount of information.

With his limit in information processing capacity there seems to be no real basis for concern about damage resulting from overloading or overtaxing man's processing mechanisms. As Miller (1963) has pointed out man is endowed with many defenses against information overload. On the contrary, the problem seems to be more one of information *underfeeding* which, when it happens at a young age, may retard the level achieved in information processing by the adult. From well documented animal studies (e.g., Riesen, 1966) it can be inferred that human beings can suffer from sensory deprivation. Templin (1957), for example, has pointed out that middle-class children generally have more learning opportunities, and, therefore, know more than children from less-educationally stimulating environments. Irwin (1960) pointed out that middle class children talk more and produce a greater variety of vowels and consonents from about 18 months on. Teachers can make a difference by getting children to verbalize more at an early age.

Not only are there great differences in the total capacity of sensory inflow between the eye and the ear and between each of them and the brain, but also there is a substantial difference in the efficiency of the sensory systems. The more pertinent information, however, is the gap between the information provided by the real world and what can be processed by sensory modalities. For example, the capacity of a television channel for transmitting information is estimated at about 5.7 $(10)^6$ bits/sec. (Bell, 1956). Television carries one million bits/sec. more than our optic nerve can handle. Estimates of channel capacity for pitch and loudness are 2.5 bits/sec. and 2.3 bits/sec. (Pollack, 1952; Garner, 1953), respectively. G. A. Miller (1951) estimated that language encoded information reaches a rate of 16 bits/sec. It also is interesting to note that the central nervous system capacity is much less than the sum of the auditory and visual modalities; therefore, its saturation can be reached by either of them. The disparity of interest here is that the central nervous system has a great storage capacity but a limited processing capacity. It also is true that information transfer is increased by increases in dimensionality, and audiovisual stimulation has more dimensionality and pathways than either audio or visual alone, consequently it probably results in more effective communication than either audio or visual alone. There are data to support this (Adams and Chambers, 1962).

Generally held to be true, but not universally accepted, is the concept of the separate storage systems for auditory and visual information (e.g., Murdock, 1960). This finding raises questions about information integration in long-term memory. Regardless of modality used, it does seem generally accepted that information goes through a multistage conversion before it can reach either short-term or long-term memory.

Summary. From this brief analysis we come to a conclusion similar to that reached from the analysis of the educational studies of media; namely, that audiovisual research efforts should be directed not at simple gross comparisons of either media or modalities, but toward the study of more specific processes and psychological variables. The processes that seem to be useful and in need of research studies are those of: (a) stimulus encoding, (b) retrieval from both short- and long-term memory, (c) both stimulus and between channel redundancy, and (d) proactive and retroactive inhibition. It would seem that interactive instructional systems should be designed to provide the means of display and control that permit manipulation of learning conditions as specified by these kinds of variables. The gross comparisons have simply used media as convenient and simplistic distinctions and the results show that the kinds of variation examined are not worth the effort. Hopefully they are a matter of history. Certainly further research support from government agencies is not indicated.

It would appear, however, that the problems posed by the functional and psychological analyses are qualitatively different. When the modes and conditions of use of media are examined in the terms of a dynamic frame of reference, the psychology of learning and information processing provide useful leads. If we are to ultimately specify the optimum design of instructional systems it will be through the use of computer-based interactive control of media and through the use of such systems to obtain the needed information.

Individual Differences and Media Effectiveness

It might be useful to examine this problem further considering some findings from studies comparing sets of correlations obtained between characteristics of learners on the one hand, and learning scores on the other hand, when different media, or instructional treatments, have been used in an effort to facilitate the learning process. Ideally, from what has been said, these studies would be richer in their implications if they had specified the differences in the treatments used in terms of the more specific kinds of variables previously described by the Tosti-Ball taxonomy. Accepting these problems with the studies, it is still possible to look at the data they have produced to get hypotheses for further research. To begin we can assume that the patterns of correlations produced by different treatments reveal differences in the characteristics of the learner that are significant in determining what the individuals have learned under the conditions imposed. If the correlational patterns differ from group to group then it could be assumed that the learning conditions differed in an interesting way. Such a result would justify an examination of their fabric for the key to their psychological differences. In other words, if the pattern of correlations obtained with each treatment were different it reveals conditions that deserve further analysis and attention.

Media as Variables—What Do Differences Mean?

We have pointed out that educational problems defined in terms of media differences produce simplistic research designs and therefore results of questionable psychological utility and questionable practical validity. Media differ from each other in a wide variety of ways some of which are intrinsic, and others of which are extrinsic and simply a matter of convention associated with their preparation or use. A film, for example, is often viewed passively and large quantities of information are presented before the viewer, or learner, who is to respond in any relevant and overt way. Typically the viewing conditions involve no overt responding, feedback, or opportunity for selective repetition. Even if a film produced a different level of learning from that produced by a self-instructional program, it would be impossible to identify the particular characteristics, or variables, that produced the difference. The programed text presents only a small amount of information at a time; it requires active overt responding by the student; it provides immediate feedback in terms of the correctness of each response, and it also can provide repetition, or even remediation, if errors occur. In short, there are many variables each of which, or a combination of which, generally differentiates programed instruction from some other approach to the design of a learning environment and thereby making comparative studies uninterpretable.

In a recently reported study (Rivers and Brudner, 1971) student characteristics were related to acquisition scores for different groups of students when the groups differed in the learning conditions to which they were exposed. These investigators reported that a cluster of verbal skill variables were related to overall. performance regardless of the media or type of task involved. In addition, a cluster of variables was found to be related to performance which was unique to each of the media involved. In the case of audiotape and videotape this second-order cluster of variables was in the visual skills area, but was more related to oral expression than to reading and test-taking skills. This suggests the hypothesis that the learners engaged in a multistage conversion or transformation process of the information as presented to them and that it involved the use of subvocal components.

They also report that in contrast to the demonstrated significance of auditory learning from taped lectures, linear and syndactic text produced acquisiton that correlated with different personality and self-interest variables. Syndactic text consists of summaries each of which is followed by a frame sequence.

Using the syndactic texts a higher level of performance was achieved for the student high in experimental personality. This is the type of individual who is more inclined to experiment in life generally and is more tolerant of inconvenience and change. It would seem that the students high in this characteristic responded to the novelty of syndactic text. In other words, students who are high in the experimental personality dimension also performed better on syndactic text.

With linear text, the first order correlations with learning scores show negative relationships for shy versus outgoing and also for exhibitionism. On the other hand, interest in the profession of librarian is positively related to performance. One inference drawn from the data is that there is a general introversion-extroversion cluster of variables involved in performance with linear text. Another is that the withdrawal type of defense mechanism achieves a higher level of performance when individuals with this personality characteristic learn under these conditions. Linear text also produces a strong relationship between rank in class and learning; this result may be more a function of motivation to study than to academic skills.

The secondary cluster of variables relating to performance resulting from computer-aided instruction (CAI) and also from a rough parallel form—audiotape and an intrinsically programed booklet—was the most difficult to identify. This may be due to the fact that these comparisons are really a composite of media and more significantly a composite of psychological variables which makes both positive and negative results difficult, if not impossible, to interpret.

Using a functional analysis (such as variations in the responses demanded of the student and in the type of remediation used when errors occurred) there appears to be a relationship between student characteristics and performance which has not been revealed by analyses involving overall performance after using media. This is the case even though the response demanded by linear text and that demanded by remediation when the syndactic text was used were different. A further indication of the value of the functional analysis of the learning environment in terms of presentation variables comes from the comparison over large segments of instruction with linear text, of a high and low response demand frequency condition. The high response demand frequency was consistently superior. A facilitation effect does appear to be produced by the insertion of a significant number of questions in the instructional material.

Although major differences were not found with the form of response required of the student, there was a trend effect revealing that the overt selected conditions were slightly better than either overt spoken or covert response.

It was found that the most efficient condition was to use moderate levels of management frequency in conjunction with high response demand frequency.

In relation to media and presentation variables the type of learning task required of the student appears to be an important factor. That is, different conditions of instruction result in differences in the rate of acquisition but not in the application of knowledge learned.

These and related data indicate that the present ways of using media as if they were useful educational variables is counterproductive. On the other hand, it appears to be productive to use variables that emerge from a psychological analysis of the learner's task. While the variables identified by Tosti and Ball are not necessarily the ultimate ones to be used, they provide, at this point in time, a rational and operational basis for the articulation of media with the processes of learning so that more personalized learning environments can be designed. It seems clear from the data that the way the teacher uses media is more important than that he uses them.

The articulation of media for instruction. The question that arises is how do we articulate media and also design environments to determine the effective principles of cognitive and motivational architecture which determine the effectiveness of a learning experience. The answer appears to be to use interactive computer-based systems. With an interactive computer-based instructional system the teacher has the capability of creating a variety of learning environments to determine which of them best facilitates the acquisition of knowledge and skills supporting the educational objectives to be achieved. While the computer languages and associated software exists for the management of media devices as a part of a learning environment, it has not been used sufficiently to provide answers to questions of environmental design that are currently needed. The architecture of learning environments is an emerging technology but more of an art at the present time. The sheer complexity of the management problems and the record-keeping justify the use of computer-based systems, and, in fact, make them necessary for the development of the technology as well as for prescriptive instruction. Among the classes of significant variables, in addition to the three described by Bruner, are the sampling and selecting of content having identified the objectives and sequencing. The ordering or sequencing of instructional presentations is clearly important. Rules for use in designing and managing an instructional program so as to produce an effective learning environment are critical. Rules for sampling content and for sequencing the subsets or units of information in combination with one another are two of the most critical determiners of the effectiveness of learning environments (Stolurow and Davis, 1965).

Another factor mentioned earlier is the catalytic role of instruction and its implications for environmental design. Teaching, in principle, should be as unobtrusive as possible; therefore, the agents of that process (e.g., teachers, machines, and schools) should be more in the background than in the foreground. The effort should be to make teaching unobtrusive. This is neither a whim or a capricious desire, but rather it is a hypothesis to be tested and it is based upon an analysis of both attitudinal and motivational dynamics of learning. Like therapy, learning should be planned so that there is a process of transference to the learner himself. In its simplest form my conception of learning as it relates to the things presented here is that it is hierarchial. At one level, a set of dynamic processes is accomplished by the learner in an organized form. Earlier, he acquires specific information processing skills that are properly cued. The results of their acquisition permit the learner to use them in many different combinations. The results of their use are revealed by the patterns of response and by the errors made during learning and transfer in such tasks as the analysis of problems and the formation of decisions. One of the critical sets of decisions to be made by the learner relates to the organization of his own instructional experiences as he solves a problem or tries to teach himself something. The design of materials for learning environments is incomplete unless it includes a set of principles for sequencing and instructional segments designed to teach them to the learner.

Sequencing. While it has long been recognized that sequencing of instruction could have an effect upon learning and the nature of the concepts learned, the demonstration of different effects has been much more difficult to achieve than expected. The interest in sequencing as a variable has been dominated by either epistomological or logical constructs. More recently the conception has shifted to the psychological aspects of the problem as a consequence of the increased interest in individualized instruction.

The development of instructional technology has contributed to the shift in emphasis from media to activity. Associated with the shift is the concern for the organization of the learner's experiences. The depth of the effect of this development is still not widely appreciated. In the older view of sequencing, or organization, the semantic aspects of the conceptual relationships within the materials presented had priority. Outlines were acceptable as a sufficient description or guide to the teacher or learner. It is now apparent that an outline is insufficient. Contributing heavily to this new perception of the requirements for a sufficient design is the computer and a distinction that is significant in programing. In mathematics typically it is sufficient to develop the formula for the solution of a problem becuase the application of the formula to the solution was accomplished by someone who could apply it and perform the steps in their necessary order and correctly. However, when the formula is written it is not sufficient for the programer. A programer needs more than a formula. He also needs the steps sequenced in a way that he can write the code which tells the machine how to process the data to get the solution. The outline and the formula are analogous and the missing element in each case is the set of steps that should be used to achieve a solution. Specification of the variables and their relationship to one another is only the beginning. An algorithm is needed; it describes a set of steps which have to be followed to solve the problem. For a complex problem, several algorithms may be needed. Each is an operational description of a solution procedure, or a processing sequence which allows a person, or a machine, to solve a specific problem. Similarily, designing learning environments—the sequencing of information for presentation to a learner—should be based upon what it is the learner is expected to do. Since different individuals might be expected to do different things with the same information it is conceivable that different instructional algorithms might be best used for each of them. The selection of an algorithm for instruction should consider the learner's needs and capabilities. The algorithm that is best for one learner may not be best for another one. It is quite clear, however, that the older and more traditional epistomological and logical orientation to the presentation of a subject matter by an author or teacher ignores the learner's problems of acquiring knowledge and associated skills for using it. This was always considered a pedogogical problem to be dealt with later by a teacher. Unfortunately this just deferred the solution of the problem and it is still with us.

At a still higher level there is the problem of determining the way to teach individuals to develop algorithms so they can produce algorithmic solutions of their own. It is interesting to note that the learner is seldom provided with algorithms for learning yet his main task in school, if not in life, is to learn and to process information. Teachers could make a tremendous difference if they were taught how to teach students to use algorithms and to develop their own algorithms. This is another way of saying that teachers should teach students how to learn. There should be built into each learning environment the algorithmic approach to problems wherever it fits and does a useful job. Where it does not, a heuristic approach should be taught. The pedogogical principle is the same, but the sequence of steps that is used is not sufficient to always generate a correct solution to the problem.

One point then is that the various conceptions of organization as algorithms and heuristics need to be developed and used in the design of learning environments whether they are live interactive systems or computer-based systems. It is obvious that sequencing has to be managed by the system, by the learner, or jointly, and by different ones of these at different times during learning. It also is obvious that the learner has to learn how to organize information in order to understand how to solve the problem. In designing learning environments the conception of sequences of information presentation to a learner should be based upon what it is that the learner must do. Since he might do different things and still solve the problem, selection of algorithms is possible and therefore different sequences can be used. Ideally the selection relates to the learner's needs. The algorithm that is best for one may not be best for another. The more traditional logical and epistomological orientations to presentation of a subject matter by an author or teacher typically ignores this problem. It is a pedogogical issue that is deferred. Consequently learning materials are more analagous to the equation from the learner's point of view. He is not provided with a basis for learning useful algorithms yet his main task is to process materials.

My point here is that an important psychological problem in the design of learning environments whether as interactive, live systems of transactional instruction or as computer-based systems, the proper conception of sequencing, is critical. It is obvious that sequencing has to be managed by the system, by the learner, or jointly. It also is obvious that the learner has to learn to organize information. Whichever one of these management approaches is used, it has to be understood or it cannot be used efficiently. Therefore, an important area to teach algorithms and heuristics of research is necessary. The explication of the effective rules of sequencing and the conditions that should be considered in determining their use are a critical set of problems to be solved if we are to properly answer the question of how teachers make a difference.

Many early studies of sequencing using programed instruction materials defined the research problem in terms of a comparison of logical with a scrambled order of presentation (e.g., Gavurin and Donahue, 1961) found adult learners reached a criterion more rapidly with logically sequenced materials; Roe, Case and Roe (1962) did not find that adults learned more rapidly from the logical sequence; Hamilton (1964) found a complex interaction effect with sequence, with response factors and with reinforcement conditions. The logical sequence typically referred to the ordering of information according to an expert in the subject-matter who was already knowledgeable. Not only is the "logical" sequence not the only logical one possible, but generally experts would not agree that the use of a particular one was optimal. While the previously referred to studies do not produce a consistent picture, there is evidence that sequence can make a difference in learning when content is controlled (e.g., Detambel and Stolurow, 1956; Anderson and Guthrie, 1966; Frase, 1970) even though the materials were not hierarchically structured. The rules that make a difference to students in determining their rate of learning are general with respect to materials (e.g., rules 1 and 2 of a synchrony). However, they are very specific with respect to the critical operations performed by the teacher or teaching system to be effective.

Sequencing also can be put under student control. When left to their own devices Mager (1964) found, for example, in teaching about electronics that students asked questions in a very different order from that used by authors of materials designed to teach such a course. Students often began with questions about vacuum tubes whereas an examination of eight outlines of basic electronic courses showed that they begin either with the topic of magnetism or with electron theory. The "logic" of the partially informed student can be quite different from the logic of the informed author or instructor. Teachers can make a difference by sequencing information to meet the needs of students. Since Mager did not compare the performance of students under the learner-generated sequence with that produced by the author-generated sequence, we do not know what differences in knowledge would result from the two different sequencing rules. Nor do we know whether learning, retention or transfer, or all of them would be affected in the same way. This kind of research should be done.

In another, Grubb (1968, 1969) presented students with a map of a statistics course on a CRT terminal as the first display. This map consisted of a series of inter-connected boxes inscribed with topics, or concepts, that were accessible to the student. This made it possible for the student to route himself through the course. Grubb (1969) reported a comparison of the learner-controlled condition with a linear version for which the content was identical. He found that students who construct their own pathway scored higher on their post-test than their counterparts, who did not. This finding identifies a psychological problem for interactive systems. The software should permit the programing of courses that allow students to generate their own paths and the system should be capable of generating "audit trails" that can be studied to determine the nature of the intuitive rules they use. Ideally the language used to program instructional materials should make it as easy for the author to develop his course for learner-controlled use as for teacher-controlled use.

Learning and problem solving are divisible into phases for the student. His perception of the task probably has a different grain from that of the teacher or observer of his performance. One common element from both points of view is testing. Usually a sequence includes a cycle involving the formulation of a testing procedure. A trial is the operation of a testing procedure, and a comparison of the results of the test with some criterion. This process has been variously described as "trial-and-error," "means-end testing," "trial-and-check," "discrepancy reduction," "test-operate-test-exit (TOTE)," "hypothesis testing," and is often characterized by a hierarchical organization. Knowledge of results has to be timed or scheduled so that it is provided when the learner completes a unit in the hierarchy.

Most learning starts with little structure so that it is clear to the learner that his response is correct only if knowledge of results is given. However, the relationship of one unit in an information sequence to the others is not revealed by knowledge of results given for correct, or incorrect response, to an individual unit. The teacher's task is to define the overall problem for the learner in terms he can use while he is learning. One approach is described by Polya (1957). He emphasizes for mathematics the necessity of determining as a first step whether it is "a problem to solve," or "a problem to define." Instruction can uniquely provide information to the learner about the higher order relevance of the current efforts he is making. In time, the learner must develop techniques for obtaining specific kinds of higher-order corrective information on his own. If the learner is to acquire this ability it is necessary for him to learn to recognize his limitations such as when he does not comprehend or understand something. To be effective at this he has to interrelate elements of information, transform them into questions which he tries to answer and perform similar activities.

The ability of problem solvers to use information to correct their erroneous hypothesis is known to vary as a function of their current state or set and it is well established that corrective information is least useful when the learner is highly motivated or anxious (e.g., Postman and Bruner, 1948). Further illustration of this was provided by Dunker (1945) who used the term "functional fixedness." It refers to using corrective information exclusively for the evaluation of a single hypothesis that happens to be wrong. An illustration is the use of a hammer for pounding when it is needed in the present problem for a weight as a bob of a pendulum. Subjects reveal remarkable intractability when motivation is high. It has been demonstrated that high drive and anxiety make learners more prone to "functional fixedness," or "persistent non-adaptive" behavior (Hamilton, 1932). When persistence is observed the immediate goal of the instruction is to terminate the state or reduce anxiety. Teaching under these conditions verges on therapy. The sequence in gross outline is clear. First treat the affective problem and then the cognitive one.

Another problem in sequencing occurs when student errors occur and must be corrected. The first problem is diagnosis. The nature of the error has to be determined. Often more information is needed. To identify the nature of errors it is necessary to have a taxonomy. Research is needed and teachers need to develop a classification system of errors which identifies the remediation procedure to be used in eliminating it. To both prevent and remediate errors information has to be coded to fit the way in which the learner is formulating the problem at the time the error is corrected. The translatability of errors is therefore critical and the instructional system has to be capable of it. A remedial capability also is needed. An extensive repertoire is needed. One instructional rule that is useful is to present corrective information either at the same or at a lower level of reading difficulty (simpler verbal encoding) than the material originally presented. For example, if the material was verbal and at an eighth grade "readability" level when the error occurred, the correction might be written at a sixth grade level of readability. This is a rule that makes use of variation in the "mode of representation" (see Bruner, 1964). Associated with this is the use of illustrative material. Since negative instances are relatively useless, they should be avoided by a teacher who wants to be effective.

Corrective information should be presented at a level below the maximum at which the student is capable of working. The student's information processing capacities should not be exceeded during either instruction or remediation. This is a rule that makes use of Bruner's concept.

Instruction is a temporary state, and the object is to make the learner, or problem solver, as self-sufficient as possible. Therefore, he should not only be corrected but also given the opportunity to learn how to correct himself. The form of the correction used at any point in time should be chosen so the learner can take over the corrective function himself.

Some Hypotheses About the Design of Learning Environments

The architecture of learning environments is probably best defined in terms of a set of rules regulating the nature and flow of the student's experiences. In contrast with physical architecture, the design of learning environments is based upon inferences regarding mental structures and functions which appear to be influenced by the nature and organization of experience during learning (e.g., Luchins, 1942). The principles of design have to be based upon inferences regarding the effects produced by observable manipulations of the learning environment.

Another factor of importance is determining the way in which mental processes come into play as a result of specific kinds of experience. For example, the encoding of one's experiences during learning (e.g., Carmeichal, Hogan and Walter, 1932; Karwaski, Gramlick and Arnott, 1944) influences the nature of the reproduction of the experience. Often the learner's, or problem solver's, task is to recode the information or to transform it. One purpose is to convert to a form that cues meaningful and relevant responses. For example, many individuals solve problems by formulating a statement about their experience. Once having coded the experience verbally the description can cue responses. People also tend to focus their attention on variables that change. With discrete display variables—for example, visually presented words—attention is focused on the covariation of a feature of a display on the one hand and a qualitative, or quantitative, change in the response that is correct on the other hand. If one or more features of a display changes from one experience to the next, the individual is likely to encode this change in the form of a statement. His statement is likely to pertain to the statement and it is likely to be related to a statement made about the response that is made. When the response is correct an association is formed between the two statements. When the response is incorrect, the relationship is not formed between the two statements but one is made with the display and knowledge that it is incorrect. Individuals are seldom exhaustive in their analyses of their errors to consider all logical possibilities; consequently, the difficulty of a learning task varies greatly depending upon the organization of the set of problems used to exemplify the concept being taught. This makes the rate of learning of a relationship highly dependent upon the sequence of instances of application.

Studies comparing the effects of logical versus random or scrambled sequence in programed instruction, for example, have failed to provide indications of what is critical in sequencing, under what kinds of conditions, and for what kinds of materials. What appears to be lacking is an objective means by which materials can be sequenced and the effects studied. Some objective sequencing rules have been developed (e.g., Detambel & Stolurow, 1956) and applied to concept learning (Anderson & Guthrie, 1966). These studies were an attempt to apply objective sequencing rules to instructional materials and to determine their effects on learning, errors made, and time required to complete the materials.

To apply these rules it is necessary to analyze the instructional materials. This is to determine what is relevant and irrelevant for a particular task or set of objectives. This analysis is similar to some undertaken in concept identification studies (e.g., Kurtz & Hovand, 1956) in which concept instances can consist of multi-leveled stimulus dimensions, some of which are relevant for the classification of the concept instance, and some of which are irrelevant.

The following are rules that have been applied to successive presentations of stimulus materials: Rule 1, if one or more relevant stimuli *change* then as many as possible of the irrelevant stimuli should be held *constant*; Rule 2, if an irrelevant stimulus *changes* then as many as possible relevant stimuli should be held *constant* Two other logical possibilities exist: Rule 3, in which both relevant and irrelevant stimuli change on successive presentations; and Rule 4, in which neither relevant nor irrelevant stimuli change. In the case of Rule 4, the same material would be presented repeatedly. Two unpublished studies add to the data previously reported by

Stolurow, (1956), and verified by Anderson and Guthrie, (1969). In one (K. Stolurow, 1971) materials consisted of programed instruction booklets in which second-grade students were taught to analyze sentences for their nouns, verbs, and adjectives. Each of the 107 students was assigned to one of three instructional groups. In each group materials were sequenced according to either Rule 1, 2, or 3. Rule 4 was not used. The dependent variables were as follows: time, errors in program, and post-test errors. Data were analyzed using a multiple analysis of co-variance with MA and pretest errors as concomitant variables.

In the other (Glick, 1971) 54 college students participated in a CAI branching program on microeconomics designed to teach such concepts as capital, short-run, and isoquants. Students were assigned to one of two treatment groups and the following measures were obtained: errors, time, post-test performance, and retention. A multiple analysis of variance and step-wise regresssion analysis were used in the data analysis.

The principle results of the studies are as follows: (a) without changing content it is possible merely by manipulating the sequence of frames within a program to significantly improve the efficiency of the program, and Rules 1 and 2 are valid ones for accomplishing that manipulation; and (b) within each rule generated sequence, achievement appears to be differentially affected by individual difference variables such as math and verbal aptitude. Therefore, the rule used determines which kinds of students will do best.

These and many other studies of sequencing in which the variables studied relate to psychological factors have shown that individuals differ in the way they process information while learning. Other related data indicate that individuals also differ in the way they retrieve information at a later time to either recall it, or to use it.

Four Basic Instructional Concepts

Four basic concepts used in transactional instruction are focused on student needs: (a) remediation, (b) inducement, (c) capitalization, and (d) compensation.

The *remedial* concept says that when you know what specific capability is missing in the learner you identify its location in the conceptional structure of the set of things to be taught and make up for it for each student who is deficient.

The *inducement* concept says that when you want to arouse thought processes, manipulate uncertainties along either, or both of two dimensions - *amount* and *saliency*. It has been reported that less uncertainty, but more saliency is best for learners who are conceptually simple and the reverse arrangement is best for those who are conceptually complex (Salomon, 1971).

The *capitalization* concept of pedagogy says that you use strategies of instruction that differ mainly in their form, or mode, rather than in their content. It says, give the learner that content he needs or wants in a form which capitalizes on his best capabilities.

The *compensatory* concept says that you use strategies in presenting content that compensates for the learner's deficiencies identified through diagnostic tests. Information, techniques, images, or formulations which a learner cannot provide for himself, are provided for him. This differs from remediation in that it anticipates problems. Generally this is more efficient than remediation which follows initial, but unsuccessful, learning.

Use of Computer-based Instructional Systems

In order for concepts of instruction such as the four just mentioned to become explicit and operationally useful on an interactive computer-based system, they must be explicit; they must be reduced to an operational form. Once that is done, then the systems software can be used to implement them in a particular computer environment. Then a subject matter has to be selected so that the *teachware* used by students can be developed. While computer-based instructional programs have been prepared by hand, one important emerging application of computers is to generate programs from a set of primitive processing routines and course files. Another is to use programs to produce processing skeletons to which there is later attached the proper content of a course. Both of these use the computer to assist the author but the latter is a two-stage process and the former uses the author for one stage.

It also is useful to distinguish between management approaches to instruction. Branching and contingency management are the two basic conceptions in vogue today. They can be used in either interventive or transactional instruction. In *branching*, the decision rule is built into the program for each of a set of specified options regarding the next event in the instructional experience. Teaching is treated as a decision process that is specified in detail and in advance. Contingency management also uses rules but it produces response-organized instructional environments. It also can use alternative sets of decision rules which themselves are hierarchically organized and contingent upon the recorded responses, and more often than not, a response history pertaining to the decision to be made.

Three classes of variables appear to be involved in developing contingencies: (a) *who* is being taught, (b) *what* is critical, and (c) *how* the teaching is to be done. Some examples of contingency rules are the following (Stolurow, 1969).

1. If the child's IQ is between 60 and 80, and he is learning to read isolated words, then it is critical to require drill and practice in which a high degree of overlearning is provided by initially using prompting, but briefly, followed by a larger confirmation series (Stolurow, 1964).

2. If an American student is high in aggression and makes incorrect responses in learning logic, then in tutorial instruction when he performs incorrectly, evaluate his responses when you tell him he is wrong; when he makes correct responses simply tell him he is correct without evaluating his response (Frase, 1963; Parisi, 1965).

3. If a student with high mathematical aptitude begins to respond more slowly (longer latency) as he works out the solution to problems that are equivalent in difficulty, then give him additional problem-solving practice at that level of difficulty but shorten the time he is allowed for solution.

The 'if" statement in each example contains a particular student characteristic. In the first example it was IQ; in the second personality (aggression) and a cultural index (American); in the third a specific aptitude (mathematical). Each rule also specifies the critical element of the instructional material or experience—reading of isolated words, logic and correctness of response; the third—speed of problem solution in mathematics. Each "then" statement includes a specific prescription for instruction—high degree of overlearning and confirmation, used or not, of evaluative feedback, and shortening of time allowed for problem solution.

Contingency statements of these kinds represent the building blocks of instructional strategies. The programing language of a CAI system should be able to accommodate sets of contingency rules, and the more conveniently it does, the easier it is to write tutorial CAI programs. The psychological research problem is to specify as many of these rules as are needed to make on-line decision processes efficient. The realization of this need has taken the form of studies designed to identify aptitude-teaching interactions, or ATI's.

Computer languages are being developed to implement these concepts. However, it also is interesting that more functional uses are being made of what appeared to be structured languages, like course writer and CAILAN. These languages have been used to accommodate the more flexible programing concepts for "unfolding" programs as required in transactional instruction. Russky, for example, is maximally unstructured: the student chooses his own learning path and is never forced to complete an item or unit. The course has no sequential pedagogical units. The only directing forces which are applied to the learner are noncoercive suggestions in the form of drills (optional), "advertisements" which remind the student of the options open to him, and periodic "mystery questions" which help the student to check his understanding of the material. The "analyze" routine is the heart of the course; its function is to recognize the student's questions and commands, sending him the appropriate information from the various blocks of information shown (included, if the student asks is the complete content of any block: a list of the prefixes of Russian, a complete printout of the grammar, etc.), and to send him mystery questions and advertisements from time to time.

Summary

The interventive and transactional conceptions of instruction while not incompatible operationally have their separate proponents and are differentially used in curriculum development. They represent ways in which teachers as well as CAI programs could relate to subject matter on the one hand, and students on the other hand. While the differences are clear, their relative effectiveness for particular kinds of objectives or people is by no means established. Clearly the problems raised require empirical research, not debate, for their solution.

The lack of knowledge about teaching stands in sharp contrast with existant knowledge about teachers and schools. Teaching is clearly more an art than a science of scientifically based technology. This makes teacher training inefficient, idiosyncratic, and subject to great swings as fads appear and disappear.

The large numbers of teachers required to keep school according to the way it is now conducted has two serious implications given the current state of the knowledge about teaching. The selection of teachers is critical but the absence of valid criteria makes the process subjective. Performance evaluation is highly variable primarily due to lack of accepted validated criteria.

Teaching is very labor intensive and the cost of education by the traditional means is exceeding the level people are willing to pay for it. The demand has typically exceeded the supply of high quality teachers who could use transactional approaches. This means that sooner or later technological aids wil be used to reduce the coverage cost per capita. Technology can relate to both instructional and research needs and it can provide the means by which a symbiotic relationship could be made a reality.

There are a number of needs to be met if teachers are to become more effective in the educational process. Teachers not only need human but also technological aids to make instruction less labor intensive. They need particular kinds of devices, not just anything called an aid, if they are to be helped in their work. They also need a type of instruction, namely that supportive of transactional instruction, if they are to promote in their students the problem solving skills and abilities.

Given the primitive state of knowledge about thinking and the ways in which these skills are learned, the problems today are to develop the concepts, methodology, and technology needed to design and construct learning environments that elicit and cultivate information processing and problem solving skills. Pertinent behavioral and social science research needs to be translated into operational principles for use in program synthesis. They also should be used to generate provisional programs for both use and testing. Teachers and authors need to learn the psychology of cognitive processes and research in this area needs to be developed with an educational focus. By also developing programs and designing research and evaluation studies teachers could learn about their subject matter in psychological as well as epistomological terms.

The current conceptions of the instructional process as the creation of learning environments or rooms for thought is a new perception of an old reality. The measure of education is not knowledge given but information processing and problem solving repertoires acquired. It is more effective to organize learning experiences in terms of inquiry and problem solving involving analysis, synthesis, and inference than in terms of what we know about the subject matter. It is more important for students to learn-how-to-learn and to want to learn than it is to teach particular facts or content whenever the former has a longer "half life" than the latter. Dispassionate inquiry and scientific neutrality while the attitudinal core of the scholar needs to be complimented in our curriculum with a sensitivity to values and cultural needs to produce effective members of society. In short, it is more important to structure learning environments that embrace content and values while developing thinking rather than force feeding large amounts of information for unidentified and often unrealized purposes.

Research on teaching in the classroom is inefficient not only because it is not cost effective, but also because the basic questions of process have been avoided. For example, trait analysis of teachers as a research question has preempted studies dealing with the functions of teachers, their strategies of teaching. Classroom studies are plagued with many unknown and uncontrolled variables. Teachers are frequently treated in a research design as if they were an educational variable or an instructional variable. This is too gross and unreliable a way of identifying critical variables. Teachers are assumed to perform consistently (reliably) and as expected (validly), although the data do not support this view. Teaching is not replicable with precision; no two performances of the same teacher on the same materials are alike. Studies need to sample teachers and students to provide generalizable results. It is difficult, if not impossible, to randomly assign teachers to strategies or types of teachings. Therefore, the teacher and the type of teaching he does are confounded in most research.

To further complicate the problems the prior history of research on learning and the vigor and rigor with which it has been pursued have resulted in learning theory eclipsing a rational analysis of instruction and the parallel development of instructional theory. Learning theories are not a sufficient basis for guiding the learning process in school subjects, nor are developmental theories. The conceptions of instruction as interventions and transactions requires a consideration of motivation, reinforcement, contingency management, and organizational problems. In addition, it is necessary to relate individual differences to methods or strategies of teaching especially if instruction is to be prescriptive and individualized.

References

Adams, J. A. and Chambers, R. W., Response to Simultaneous Stimulus of Two Sense Modalities. *Journal of Experimental Psychology*, 1962, 63, pp. 198-206.

Barrow, L. C. and Westley, B. H., Exploring the News: An Experiment on the Relative Effectiveness of Radio and TV Versions of a Children's News Program. *A. V. Communications Review,* 1959, 7, 14-23.

Bell, D. A., *Information Theory and its Engineering Application*, New York: Pitman, 1956.

Berlyne, D. E., *Conflict, Arousal and Curiosity.* New York: McGraw-Hill, 1960.

Bloom, B. S., et al.: (Ed.'s.) *Taxonomy of Educational Objectives. Handbook I: Cognitive Domain*. New York: McKay, 1956.

Broadbent, D. E., *Perception and Communication*. London: Pergamon Press, 1958.

Bruner, J. S., *The Process of Education*. Cambridge, Mass: Harvard University Press, 1960.

Bruner, J. S., Some Theorems on Instruction Illustrated with Reference to Mathematics. In E. R. Hilgard (Ed.), *Theories of Learning and Instruction*. Chicago: University of Chicago Press, 1964, pp. 306-335.

Bruner, J. S., *Toward the Theory of Instruction*. Cambridge, Mass: Harvard University Press, 1966.

Bryan, G. L., Computers and Education. *Computers and Automation*. 1969, 18, pp. 1-4.

Campeau, P. L., Level of Anxiety and Presence or Absence of Feedback in Programmed Instruction. Palo Alto, Calif: American Institutes for Research, Feb., 1966 (mimeo).

Carbonell, J. R., A. I. In CAI: Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*. 1970, mms-11, 4, pp. 190-202.

Ellis, H. C., Research Design in Programmed Learning. *Proceedings of the National Society for Programmed Instruction*. August, 1962, pp. 7-8.

Feldman, D. and Sears, P., Effects of Computer-Assisted Instruction on Children's Behavior. *Educational Technology*, March, 1970, 10, pp. 11-14.

Feurzeig, W., et al. Programming-Languages as a Conceptual Framework for Teaching Mathematics. *Interface*, 1970, 4, pp. 13-17.

Frick, F. C., Some Perceptual Problems from the Point of View of Information Theory. N. B. McMillan, et al. (Ed.'s.) *Current Trends in Information Theory*. Pittsburgh: University of Pittsburgh Press, 1953.

Furey, M. Z., Results of Summative Analysis of a Multi-Media Economics Course. Paper read at 1971 AERA annual meeting. New York, February 5, 1971.

Garner, W. R., An Information Analysis of Absolute Judgment of Loudness. *Journal of Experimental Psychology*, 1953, 46, pp. 373-380.

Garner, W. R., *Uncertainty and Structure as Psychological Concept*. New York: Wiley, 1962.

Grubb,R. E., Learner-Controlled Statistics. *Programmed Learning*, January, 1968, 38-42.

Halloran, J. D., *The Effects of Math Communication with Special Reference to Television*. Leicester, England: Leicester University Press, 1964.

Hess, R. D., et al, The Computer as a Socializing Agent: Some Socio-Effective Outcomes of the CAI, Stanford, Calif: Stanford University, Stanford Center for Research and Development in Teaching. Technical Report No. 13, 1970.

Hilgard, E. R. and Bower, G. H., *Theories of Learning*. New York: Appleton-Century-Crofts, 1966.

Holland, J. G., Research on Programming Variables in R. Glaser (ed.), *Teaching Machines and Programmed Learning, II: Data and Directions*. Washington, D.C.: National Educational Association, 1965, pp. 66-117.

Holmes, P. D., Jr., *Television Research in the Teacher-Learning Process*. New York: National Education Television and Radio Center, 1960.

Irwin, O. C., Language and Communication. In P. H. Mussen (Ed.). *Handbook of Research Methods in Child Development*. New York: Wiley, 1960.

Krathwohl, D. R., et al. *Taxonomy of Educational Objectives. The Classification of Educational Goals: Handbook II: Affective Domain*. New York: McKay, 1964.

Kumata, H., *An Inventory of Instructional Television Research*. Ann Arbor: Educational Television and Radio Center, 1956.

Kumata, H., A Decade of Teaching by Television. In W. Schramm (Ed.) *The Impact of Educational Television*. Urbana, Ill.: University of Illinois Press, 1960.

Manwell, T., Daugherty, W., Desch, F. and Stolurow, L. M., Toward the Interactive Teaching of Russians, Cambridge, Mass: Harvard CAI Laboratory, Technical Report No. 1, 1970.

Miller, G. A., *Language and Communication*. New York: McGraw-Hill, 1951.

Miller, G. A., The Magic Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review*, 1956, 63, pp. 81-97.

Miller, G. A., *The Psychology of Communications*. New York: Basic Books, Inc., 1967.

Miller, J. G., The Individual as an Information Processing System. In W. S. Fields and W. Abbott (Ed.'s.). *Information Storage and Neural Control*. Springfield, Ill: Charles C. Thomas, 1963, pp. 301-308.

Murdock, B. B., The Immediate Retention of Unrelated Words. *Journal of Experimental Psychology*, 1969, 60, pp. 222-234.

Osborne, J. W., Quastler, H., and Tweedell, K. S., Flash Recognition—Scale Management. Urbana, Ill.: Control System Laboratory, University of Illinois, Report No. 78, 1955.

Pollack, I., The Information of Elementary Auditory Displays. *Journal of Acoustical Society of America*, 1952, 24, pp. 745-750.

Reid, J. C. and MacLennan, D. W., *Research in Instructional Television and Films*. Washington, D.C.: U.S. DHEW, Bureau of Research. (Catalogue No. FS 5.234: 34041) 1967.

Riesen, A. H., Sensory Deprivation. In E. Stellar and J. M. Sprague: (Ed.'s.). *Progress in Physiological Psychology*, New York: Academic Press, 1966.

Rivers, L. C. and Brudner, H., Research and Evaluation in a Multi-Media Psychology/Leadership/Management Course. Paper read at 1971 AERA annual meeting. New York, February 4-7, 1971.

Rosen, B. C. and Race, F., Ethnicity and the Achievement Syndrome. *American Sociological Review*, 1959, 24, pp. 47-60.

Salomon, G., Heuristics for the Generation of Aptitude—Treatment—Interaction Hypothesis. Paper read at 1971 AERA annual meeting. New York, February 5, 1971. (mimeo).

Scramm, W. (Ed.) *Educational Television: The Next Ten Years*. Stanford, Calif.: Stanford University Press, Institute of Communication Research, 1962, pp. 52-76.

Shaftel, F. R., Crabtree, C. and Rushworth, V., Problem-Solving in the Elementary School. In R. E. Gross (Ed.). Problems Approach in the Social Studies. Washington, D.C.: National Education Association, 1969, pp. 25-47.

Skinner, B. F., *Cumulative Record*, (Revised Ed.) New York: Appleton-Century-Crofts, 1961.

Staats, A. W. and Staats, C. K., *Complex Human Behavior*. New York: Holt, Rinehart, and Winston, Inc., 1963.

Starkweather, J. A., A Common Language for a Variety of Conversational Programming Needs. In R. C. Atkinson and H. A. Wilson (ed.'s.). *Computer-Assisted Instruction, A Book of Readings*, New York: Academic Press, 1969, pp. 269-304.

Stickell, D., A Critical Review of the Methodology and Results of Research Comparing Television and Face-To-Face Instruction. Unpublished doctoral dissertation. Pennsylvania State University, 1963.

Stolurow, L. M., A Test of the Hypothesis that the Variability of Behavior in Albino Rat is an Index of Problem-Solving Behavior. Unpublished Master's Thesis, Cornell University, 1946.

Stolurow, L. M., Implications of Current Research and Future Trends in J. P. DeCecco (ed.). *Educational Technology*. New York: Holt, Rinehart, and Winston, 1964, pp. 432-446.

Stolurow, L. M., Computer-Assisted Instruction in S. M. McMurrin (ed.). *The Schools and the Challenge of Innovation*. New York: McGraw-Hill, 1969a, pp. 270-319.

Stolurow, L. M., Some Factors in the Design of Systems for Computer-Assisted Instruction, chapter in R. C. Atkinson and H. A. Wilson (ed.'s.). *Computer-Assisted Instruction, A Book of Readings*. New York: Academic Press, 1969b, pp. 65-93.

Stolurow, L. M., Computer-Aided Instruction—Theory and Practice in A. Mondadrii (ed.) *Automation and Educational Problems*. Milan, Italy: FAST, 1969c, pp. 243-278.

Stolurow, L. M. and Davis, D. J., Teaching Machines and Computer-Based Systems, chapter in R. Glazer (ed.) *Teaching Machines and Programmed Learning: II, Data and Directions*. Washington, D.C.: National Education Association, 1965, pp. 162-212.

Strodtbeck, F. L., Family Interaction, Values, and Achievement in D.C. McClelland, et al (ed.'s.) *Talent and Society: New Perspective in the Identification of Talent*. Princeton, N.J.: Van Noustrand Co. 1958.

Templin, M. C., Certain Language Skills in Children: *Their Development and Interrelatioships*. Minneapolis: University of Minnesota Press, 1957.

Tosti, D. and Ball, J. R., A Behavioral Approach to Instructional Design and Media Selection, Albuquerque, N. M.: Behavioral Systems Division, Westinghouse Learning Corp., 1969.

Table 1

SYSTEM CAPABILITY MATRIX AND DIMENSIONS OF PRESENTATION*

System Capability	Attributes of System Capability	
	Form	Frequency
Stimulus	*Stimulus Representation* Verbal-written Verbal-spoken Pictorial	*Duration* Transient-Persistant Length of time the presentation remains intact a. low b. intermediate c. high
Response	*Response Demand* Overt-written Overt-spoken Covert	*Response-Demand Frequency* Infrequent-Frequent Frequency of response required a. low or zero b. intermediate c. high
Management	*Management Form* Repetition Multilevel Multiform Error-diagnostic	*Management Frequency* Infrequent-Frequent Frequency of decision to change presentation a. low or zero b. intermediate c. high

*After Tosti and Ball, 1969

NOTES ON CONTRIBUTORS

1. N. L. Gage – is Acting Director of the Stanford Center for Research and Development in Teaching, Palo Alto, California. He is also Professor of Educational Psychology, Stanford University, and editor of the *Handbook of Research on Teaching.*

2. Ned Flanders – is Program Director II, Far West Laboratory for Educational Research and Development, Berkeley, California. Beginning September 1, 1971 he will be Acting Director of the Teacher Education Program. He was previously professor and consultant on research, University of Michigan, and is the author of *Analyzing Teaching Behavior,* 1970.

3. Philip M. Jackson – is Director of the Laboratory Schools, at the University of Chicago where he is also Professor of Education and Human Development. He is the author of *Life in Classrooms,* 1968.

4. Dan Lortie – is Associate Professor in the Department of Education at the University of Chicago, and is associated with the Midwest Administration Center there. Dr. Lortie is a sociologist who has conducted occupational studies of both medicine and law and is currently engaged in studies of teaching as an occupation and the public schools as organizations.

5. Alexander M. Mood – is former Assistant Commissioner for the National Center of Educational Statistics in the U.S. Office of Education. In that capacity he supervised the conduct of the Equality of Educational Opportunity Survey (often called the Coleman Report). Dr. Mood is currently the Director of the Public Policy Research Organization, University of California at Irvine.

6. Barak Rosenshine – is Associate Professor in the Bureau of Educational Research, University of Illinois, where he is engaged in writing reviews and conducting research in the area of teacher behavior and student achievement. He is a contributor to the recent book *Research on Teacher Education: A Symposium,* edited by B. O. Smith.

7. Lawrence Stolurow – is Director of Graduate Studies, Office of the Provost for Educational Research and Development, State University of New York at Stony Brook. From 1966 to 1971 he was Director of the Computer Assisted Instruction (CAI) Laboratory, Harvard University. Dr. Stolurow is the Author of *Teaching by Machine,* an HEW cooperative research monograph, 1961.

CONTRIBUTORS AND DISCUSSANTS

Contributors

Dr. N. L. Gage
Acting Director
Stanford Center for Research and Development in Teaching
Palo Alto, California

Dr. Ned A. Flanders
Program Director II
Teacher Education Program
Far West Laboratory for Educational Research and Development
Berkeley, California

Dr. Philip Jackson
Director
Laboratory Schools
University of Chicago

Dr. Dan C. Lortie
Associate Professor of Education
Midwest Administration Center
University of Chicago

Dr. Barak Rosenshine
Associate Professor
Bureau of Educational Research
College of Education
University of Illinois
Urbana, Illinois

Dr. Alexander M. Mood
Director, Public Policy Research Organization
University of California at Irvine
Irvine, California

Dr. Lawrence M. Stolurow
Director of Graduate Studies
Office of the Provost for Educational Research and Development
State University of New York
Stony Brook, New York

Discussants

Dr. J. Victor Baldridge
Assistant Professor of Education and Sociology
School of Education
Stanford University
Stanford, California 94305

Dr. Walter R. Borg
Director
Teacher Education Program
Far West Laboratory for Educational Research and Development
Berkeley, California

Dr. William Cooley
Co-Director
Learning Research and Development Center
Pittsburgh, Pennsylvania

Dr. James M. Cooper
Assistant Professor
School of Education
University of Massachusetts
Amherst, Massachusetts

Dr. Jimmie C. Fortune
Professor, School of Education
University of Massachusetts
Amherst, Massachusetts

Dr. Wells Hively II
Chairman, Psychological Foundations of Education
University of Minnesota
Minneapolis, Minnesota

Dr. Charles Leyba
Director, Project Maestro
School of Education
California State College
Los Angeles, California

Dr. Lawrence W. Lezotte
Assistant Director for Research
Center for Urban Affairs
Owen Graduate Center
Michigan State University
East Lansing, Michigan

Dr. Gordon L. McAndrew
Superintendent of Schools
School City of Gary
Gary, Indiana

Dr. Frederick J. McDonald
Chairman, Teacher Behavior Research Group
Educational Testing Service
Princeton, New Jersey

Miss Yvonne D. McElroy
Assistant Program Director
Woodlawn Experimental Schools District Project
Chicago, Illinois

Dr. Charles E. Scruggs
Alfred North Whitehead Fellow
Graduate School of Education
Harvard University
Cambridge, Massachusetts

Mr. David X. Spencer
Chambers Memorial Baptist Church
New York, New York

☆ U.S. Government Printing Office: 1971 O—437—572

UNITED STATES
GOVERNMENT PRINTING OFFICE
PUBLIC DOCUMENTS DEPARTMENT
WASHINGTON, D.C. 20402

OFFICIAL BUSINESS

POSTAGE AND FEES PAID
U.S. GOVERNMENT PRINTING OFFICE

U.S.MAIL

Asst. Comr., National Center for
Educational Communication
Room 3009 - FOB-6

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education/OE-58044